

California State University, San Bernardino

CSUSB ScholarWorks

Theses Digitization Project

John M. Pfau Library

2000

Some factors that effect [sic] statistical power in ANCOVA: a population study

Valerie Maria Tvedt

Follow this and additional works at: <https://scholarworks.lib.csusb.edu/etd-project>



Part of the [Applied Behavior Analysis Commons](#)

Recommended Citation

Tvedt, Valerie Maria, "Some factors that effect [sic] statistical power in ANCOVA: a population study" (2000). *Theses Digitization Project*. 1764.

<https://scholarworks.lib.csusb.edu/etd-project/1764>

This Thesis is brought to you for free and open access by the John M. Pfau Library at CSUSB ScholarWorks. It has been accepted for inclusion in Theses Digitization Project by an authorized administrator of CSUSB ScholarWorks. For more information, please contact scholarworks@csusb.edu.

SOME FACTORS THAT EFFECT STATISTICAL POWER IN ANCOVA:

A POPULATION STUDY

A Thesis

Presented to the

Faculty of

California State University,

San Bernardino

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

in

Psychology

by

Valerie Maria Tvedt

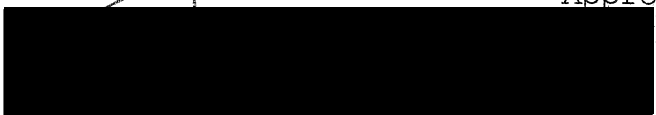
June 2000


SOME FACTORS THAT EFFECT STATISTICAL POWER IN ANCOVA: A
POPULATION STUDY


A Thesis
Presented to the
Faculty of
California State University,
San Bernardino

by
Valerie Maria Tvedt
June 2000

Approved by:


Dr. Jodie Ullman, Chair, Psychology


Dr. Jan Kottke


Dr. Kenneth Schultz

6/9/00
Date

ABSTRACT

A study into the factors that effect power in an analysis of covariance (ANCOVA) design were examined. Four factors - sample size, significance level, dependent variable-covariate correlations and homogeneity of regression - were varied in a population study. Results indicate that power increased when the dependent variable-covariate correlations increased and when sample size increased. Power also increased when a less stringent alpha level was used. Homogeneity of regression did not effect power. Implications and recommendations for the applied researcher are discussed.

ACKNOWLEDGMENTS

The author wishes to acknowledge the unending support of the following individuals: Dr. Jodie Ullman for her guidance, input and unending patience with this thesis; Dr. Jan Kottke and Dr. Kenneth Shultz for sitting on my committee and their input towards the completion of this thesis. I would also like to acknowledge Aysheh Antuono for taking on the daunting task of organizing, inputting and editing both the data and the thesis and Steve Lampman for further organizing, creating and editing all figures.

TABLE OF CONTENTS

ABSTRACT	iii
ACKNOWLEDGMENTS	iv
INTRODUCTION	1
Analysis of Variance (ANOVA)	3
Analysis of Covariance (ANCOVA)	8
Difference between ANOVA and ANCOVA	12
Factors that Effect Power	
Effect Size and Power	14
Sample Size	16
Significance Level	18
Measurement Error	19
Dependent Variable-Covariate Correlations	20
Homogeneity of Regression	22
Hypotheses	23
METHOD	24
Procedure	24
ANCOVA Design	24
Calculating the Non-Centrality Parameter	25
Models that Assume Homogeneity of Regression .	25
Models that Examine the Effects of Violation of Homogeneity of Regression	26
Variables	27

Design	28
RESULTS AND DISCUSSION	31
Sample Size	31
Significance Level	33
Dependent Variable-Covariate Correlations	36
Homogeneity of Regression	39
Overall Conclusions for the Applied Researcher	43
APPENDIX A: Design of Experiment	
Figure 1: Model of design with sample sizes (n), $\alpha = .01$, DV-covariate correlations, and IV-covariate interactions (η^2)	45-46
Figure 2: Model of design with sample sizes (n), $\alpha = .05$, DV-covariate correlations, and IV-covariate interactions (η^2)	47-48
APPENDIX B: EQS Input Files	
Figure 1: Sample program file for testing no violation of homogeneity of regression	49
Figure 2: Sample program file for testing violation of the homogeneity of regression assumption ..	50
APPENDIX C: Power Results by Sample Size, Significance Level, and Correlations	
Figure 1: Power values across DV-Covariate correlations (r) and sample sizes for $\alpha = .01$ and IV-Covariate interaction (η^2) $\leq .05$	51-52

Figure 2: Power values across DV-Covariate correlations (\mathbf{r}) and sample sizes for $\alpha = .01$ and IV-Covariate interaction (η^2) = .30	53-54
Figure 3: Power values across DV-Covariate correlations (\mathbf{r}) and sample sizes for $\alpha = .01$ and IV-Covariate interaction (η^2) = .40	55-56
Figure 4: Power values across DV-Covariate correlations (\mathbf{r}) and sample sizes for $\alpha = .01$ and IV-Covariate interaction (η^2) = .50	57-58
Figure 5: Power values across DV-Covariate correlations (\mathbf{r}) and sample sizes for $\alpha = .01$ and IV-Covariate interaction (η^2) = .70	59-60
Figure 6: Power values across DV-Covariate correlations (\mathbf{r}) and sample sizes for $\alpha = .05$ and IV-Covariate interaction (η^2) $\leq .05$	61-62
Figure 7: Power values across DV-Covariate correlations (\mathbf{r}) and sample sizes for $\alpha = .01$ and IV-Covariate interaction (η^2) = .30	63-64
Figure 8: Power values across DV-Covariate correlations (\mathbf{r}) and sample sizes for $\alpha = .01$ and IV-Covariate interaction (η^2) = .40	65-66

Figure 9: Power values across DV-Covariate correlations (\mathbf{r}) and sample sizes for $\alpha = .01$ and IV-Covariate interaction (η^2) = .50	67-68
--	-------

Figure 10: Power values across DV-Covariate correlations (\mathbf{r}) and sample sizes for $\alpha = .01$ and IV-Covariate interaction (η^2) = .70	69-70
---	-------

APPENDIX D: Power Results for Homogeneity of Regression

Figure 11: Power values across DV-Covariate correlations (\mathbf{r}) and IV-Covariate interactions (η^2) for $\alpha = .01$ and $n = 30$	71-72
--	-------

Figure 12: Power values across DV-Covariate correlations (\mathbf{r}) and IV-Covariate interactions (η^2) for $\alpha = .01$ and $n = 60$	73-74
--	-------

Figure 13: Power values across DV-Covariate correlations (\mathbf{r}) and IV-Covariate interactions (η^2) for $\alpha = .01$ and $n = 90$	75-76
--	-------

Figure 14: Power values across DV-Covariate correlations (\mathbf{r}) and IV-Covariate interactions (η^2) for $\alpha = .01$ and $n = 120$	77-78
---	-------

Figure 15: Power values across DV-Covariate correlations (\mathbf{r}) and IV-Covariate interactions (η^2) for $\alpha = .01$ and $n = 150$	79-80
---	-------

Figure 16: Power values across DV-Covariate correlations (\mathbf{r}) and IV-Covariate interactions (η^2) for $\alpha = .05$ and $n = 30$	81-82
Figure 17: Power values across DV-Covariate correlations (\mathbf{r}) and IV-Covariate interactions (η^2) for $\alpha = .05$ and $n = 60$	83-84
Figure 18: Power values across DV-Covariate correlations (\mathbf{r}) and IV-Covariate interactions (η^2) for $\alpha = .05$ and $n = 90$	85-86
Figure 19: Power values across DV-Covariate correlations (\mathbf{r}) and IV-Covariate interactions (η^2) for $\alpha = .05$ and $n = 120$	87-88
Figure 20: Power values across DV-Covariate correlations (\mathbf{r}) and IV-Covariate interactions (η^2) for $\alpha = .05$ and $n = 150$	89-90
REFERENCES	91

INTRODUCTION

The general purpose of this research was to explore statistical power in the context of analysis of covariance (ANCOVA). Using sample data, a researcher attempts to design an experiment that is sensitive enough to detect differences that might be present in the population being measured. Different designs and statistical analyses will have varying effects on power. Since power is the ability to detect a difference among treatment effects if such differences exist, it is defined as $1-\beta$, where β is the probability of making a type II error. Type II error is retaining a false null hypothesis or missing an effect that was present. We can create sensitivity or powerful analyses by using large sample sizes, by choosing treatment conditions that are expected to produce sizable effects (e.g. using no drug versus using a high dose of the same drug), utilizing a less stringent significance level ($p < .05$ versus $p < .01$) and by reducing the uncontrolled variability within the study (e.g. using a covariate within the statistical design) (Keppel, 1991).

The power of an experiment is determined by the interaction of three factors - significance level α , the

magnitude of the treatment effects, and sample size, n (Keppel, 1991). Kraemer's study (as cited in Keppel, 1991) stated that the following factors influence power: 1) increasingly larger sample sizes are needed to increase power by a fixed amount; 2) relatively small expected effect sizes lead to reduced power, and 3) adopting a less stringent significance level leads to increased power. For example, what if a researcher wants to increase the power of her experiment from .50 to .80? According to the first factor, the researcher would potentially need to obtain a greater number of participants for the experiment in order to increase the power by that interval. However, as seen in the above stated rules, just increasing sample size alone may not be the complete answer nor the best one. The researcher could also try obtaining a greater effect size by increasing the intensity of the treatment condition and/or by reducing the error within the design. Another option available is the researcher could select a less stringent criteria level, such as deciding to set her significance level to $p < .05$ instead of $p < .01$.

Within the context of the experiment as a whole, power is related to two considerations: the overall design of an experiment (whether to use a completely randomized between-

subjects design versus a completely within-subjects design) and the statistics used to analyze the data. Because consideration of the overall design is beyond the scope of this experiment, only the model of analysis of covariance (ANCOVA) will be examined. For simplification both in discussion and computations, only equal n will be considered. Yet, before addressing ANCOVA, the basic analysis of variance (ANOVA) model must be explicated.

Analysis of Variance (ANOVA)

ANOVA is a statistical method that measures the ratio between the treatment variance and the error variance. Stated another way, it is the ratio of the between-group variance within the experiment to the within-group variance.

ANOVA has its basis in the General Linear Model which demonstrates that within a single-factor experimental design, there exists three elements. These elements are best illustrated in the following equation:

$$Y_{ij} = \mu_T + \alpha_i + \varepsilon_{ij} \quad \text{Equation 1}$$

where Y_{ij} = one observation in any of the treatment groups

μ_T = the grand mean of the treatment populations

α_i = is the treatment effect for a condition, and

ϵ_{ij} = the experimental error.

The linear model concisely expresses all the factors influencing the results of any given condition within the experimental manipulation (Keppel, 1991). ANOVA has three basic assumptions that are intrinsic to its function. The first is the assumption of homogeneity of variance, which states that the variance within the groups being tested is approximately equal across the groups in the design (Keppel, 1991). The second assumption is that of normality. Normality states that the individual treatment populations are normally distributed (Keppel, 1991). The final assumption for the ANOVA design is the independence of errors. This assumption states that any given score has no influence on any other scores either within the treatment group or across the groups of treatment.

There is also another aspect within ANOVA that needs to be addressed -- the relationship found within the F ratio itself. From statistical theory we know that the F ratio is comprised of the ratio of treatment variance to error variance, symbolically stated as:

$$\frac{\sigma_A + \sigma_{S/A}}{\sigma_{S/A}}$$

Equation 2

where σ_A = an estimate of variance of the treatment effect

$\sigma_{S/A}$ = an estimate of the error variance

It is important to note that Equation 2 contains estimates of population values. Since the population values are not available for research, the F ratio can be examined through expected values. The expected values for $MS_{S/A}$, $E(MS_{S/A})$, known as the within-groups mean square, is obtained through repeated random samplings that would produce a sampling distribution with a mean variance found in the population (Keppel, 1991). When stated symbolically, the relationship becomes:

$$E(MS_{S/A}) = \sigma^2_{\text{error}}$$

Equation 3

with $E(MS_{S/A})$ = the expected value for $MS_{S/A}$ and

σ^2_{error} = the population error variance.

The expected value of MS_A , $E(MS_A)$, represents a combination of both the treatment component and error variance and is symbolized as follows:

$$\frac{E(MS_A) = \sigma^2_{\text{error}} + n \sum (\alpha_i)^2}{a-1}$$

Equation 4

with σ^2_{error} = the population error variance

n = number of equal observations in each cell
contributing to the estimate of each
treatment mean

$\Sigma(\alpha_I)^2$ = the sum of the treatment effects and

$a-1$ = the number of levels of the independent
variable minus 1 for correction.

When the two components are combined into the F ratio, it becomes the following:

$$\frac{E(MS_A)}{E(MS_{S/A})} = \frac{\sigma^2_{\text{error}} + n \Sigma(\alpha_I)^2 / a - 1}{\sigma^2_{\text{error}}} \quad \text{Equation 5}$$

By examining the ratio, the utility of the ANOVA design becomes apparent. By dividing the treatment by the error, it is expected that the effect seen is that of the true treatment effect. With closer examination of the ratio, the statistical power of the model can be identified. Note that the error term is present in both the numerator and the denominator of the ratio. One source of statistical power for the model originates from the size of the error term in the denominator. If the error term is small, the F ratio will be larger. With this larger ratio, there is

greater opportunity that the design will detect a treatment effect on the dependent variable. This in turn can lead to greater power.

An example of how expected values can be used by the applied researcher is as follows: suppose a researcher was interested in the effects of sleep deprivation on math performance. The researcher hypothesized that the more sleep deprived an individual was, the poorer that individual's math performance on a timed math test would be. The researcher divides the participants, a sample of the population she is interested in, into three levels of sleep deprivation, administers the three levels of treatment, and then the participant's math performance is tested using a timed math test. Once the data is gathered, the researcher performs an ANOVA, which utilizes the expected values, to see if there are significant differences between treatment conditions. ANOVA would yield a ratio of expected values that would be compared with its corresponding critical ratio. At this point, the null hypothesis (that the amount of sleep deprivation has no effect on an individual's math performance) would either be retained or rejected based upon a sample of the population and the expected values.

Analysis of Covariance (ANCOVA)

The second statistical model that will be considered is analysis of covariance (ANCOVA). ANCOVA "is an extension of analysis of variance in which main effects and interactions of IVs are assessed after DV scores are adjusted for differences associated with one or more covariates" (Tabachnick & Fidell, 1996, p, 321).

Like ANOVA, analysis of covariance has the General Linear Model as its basis but with one extra factor -- the addition of a covariate. Thus, the symbolic representation of this revised linear model would be as follows:

$$Y_{ij} = \mu_T + \alpha_i + \varepsilon_{ij} - \varepsilon_{cov} \quad \text{Equation 6}$$

where Y_{ij} = one observation in any of the treatment groups

μ_T = the grand mean of the treatment populations,

α_i = the treatment effect for a condition,

ε_{ij} = the experimental error, and

ε_{cov} = error removed by the correlation between the dependent variable and the covariate.

ANCOVA has five (5) assumptions that are implicit to its design. The first three are identical to the ANOVA design (homogeneity of variance, normality, and independence of errors). The last two (2) are homogeneity of regression and the assumption of linear regression (Keppel, 1991). Homogeneity of regression states that the within-group regression coefficient (the average of the regression coefficients of the treatment groups) is equal. The assumption of linear regression states the "the deviations from regression-that is, the residual scores-are normally and independently distributed in the population, with means of zero and homogeneous variances" (p. 316). This assumption states that the true regression is linear and that a violation of this assumption would suggest that the adjustment made is not as beneficial as the same adjustment made with a true linear regression.

The new statement of this model is that for any given observation of the dependent variable Y_{ij} , there are four factors affecting that observation. The first one is the grand mean of treatment populations, represented as μ_T . Added to this grand mean is the treatment effect, represented as α_i . The first error term, ϵ_{ij} , is the

unaccountable error within the measurement of the observation. The last term, ε_{cov} , represents that error that is being removed from the equation due to the relationship between the dependent variable and the covariate. This last term within the linear model is the key difference between ANOVA and ANCOVA. Recall that the linear model for analysis of variance is:

$$Y_{ij} = \mu_T + \alpha_i + \varepsilon_{ij} \quad \text{Equation 1}$$

Note that with ANOVA, there is no specification as to how to further reduce the error. This effect that error has on ANOVA can be further demonstrated when the ratio of expected values is examined. Recall again, the F ratio for ANOVA in terms of expected values is:

$$\frac{E(MS_A)}{E(MS_{S/A})} = \frac{\sigma^2_{\text{error}} + n \sum (\alpha_i)^2 / a - 1}{\sigma^2_{\text{error}}} \quad \text{Equation 5}$$

As can be seen within the ratio, error is present in both the numerator and the denominator. Logic and mathematics dictates that the smaller the denominator in any fraction, the larger the number once that fraction is converted into decimals. The same logic applies to Equation 5 in that the smaller the error term, (the denominator) the larger the F ratio, the greater the power.

The ANCOVA model, in contrast, attempts to reduce the error term by correcting the observations of the dependent variable after the effect of a covariate has been removed. This overall correction would be symbolically represented as

$$\frac{E(MS_A)}{E(MS_{S/A})} = \frac{\sigma^2_{\text{error}} + n (\sum (\alpha_i)^2 / a - 1) - (\sigma'_y)^2}{\sigma^2_{\text{error}} - (\sigma'_y)^2} \quad \text{Equation 6}$$

where $\sigma'_y = \sigma_y \sqrt{1 - r^2}$, the within-population of the adjusted Y' .

By adding σ'_y to the equation, a statistical correction has been made to the error variance such that it has been reduced. Reducing the error variance allows for more of the treatment effect to be observed without that effect being clouded by error. According to Maxwell, Delaney and Dill (1984), this reduction of error (by the use of a covariate) leads to a more precise estimate of the treatment effect and increased statistical power. Thus, the ANCOVA design is often more powerful because of its increased sensitivity in detecting a statistically significant difference.

Although power is increased by using ANCOVA instead of ANOVA, at least two important questions remain unanswered:

1) how large should the correlation between the dependent variable - covariate be in order to maximize power? and 2) what is the effect of violation of the homogeneity of regression assumption on power?

Difference between ANOVA and ANCOVA

Before addressing these questions, a detailed comparison of ANOVA and ANCOVA is warranted. There are some considerations that have to be made when using ANCOVA. In contrast to ANOVA (which requires that there be an independent variable and a dependent variable), the researcher using an ANCOVA design relies on the addition of a covariate to increase detection of a true difference. Recall that the formula for the correlation within ANCOVA is:

$$r = \frac{\text{cov}(DV \& cov)}{\sqrt{(\text{var } DV)(\text{var } cov)}} \quad \text{Equation 7}$$

This formula, when transformed into population parameters, becomes ρ within the following expected values formula:

$$\frac{E(MS_A)}{E(MS_{S/A})} = \frac{\sigma^2_{\text{error}} \sqrt{1-\rho} + n (\sum (\alpha_i)^2 / a - 1)}{\sigma^2_{\text{error}} \sqrt{1-\rho}} \quad \text{Equation 8}$$

where ρ = population correlation coefficient

It is in Equation 9 that the conceptual key difference between ANOVA and ANCOVA can be seen. Because ρ is defined as the population correlation coefficient, when $\rho = 0$, the expected values equation yields an ANOVA design, because there was no correction made to the error term. As ρ increases from 0, the ANCOVA design emerges. Notice that as ρ increases, the correction made to the error term increases. If $\rho = 1$, there is no error because $\sqrt{1-\rho} = \sqrt{1-1} = 0$, implying that all the error has been accounted for with the correlation relationship. Further, if the error term is adjusted by the correlation, it becomes smaller and this smaller error gives an increase in power. Cohen (1988) stated the relationship more clearly: "the ANCOVA design yields greater power, in general, than ANOVA because the within-population σ of the adjusted Y' variable will be smaller than σ of the unadjusted Y variable" (p.380).

To put these differences between ANOVA and ANCOVA into perspective, recall the example in which the researcher was testing the effects of sleep deprivation on math performance. Suppose the researcher wished to add a covariate to the experiment in an effort to reduce the overall error term, such as math proficiency. The

researcher could select subjects based upon their level of proficiency in math, hypothesizing that there is a high correlation between the covariate (math proficiency) and the dependent variable (math performance), and that such a correlation will help account for and remove some of the error from the treatment effect (amount of sleep deprivation).

Although both ANOVA and ANCOVA have been discussed, only ANCOVA and the specific factors that can effect power in ANCOVA will be examined.

Factors that Effect Power

Effect Size and Power

ANCOVA can determine if there was a statistically significant treatment effect as well as indicate a magnitude of that effect. However, of the total variance that can is associated with the dependent variable, how much of that variance can be explained by the independent variable? In order to answer this new question and facilitate the power of the experiment, the researcher can calculate what is called an effect size. An effect size is a treatment magnitude index that refers to the proportion of variation "explained" or "accounted for" by the treatment manipulation in an experiment (Keppel, p.65).

A measure of treatment effect that is popular is eta squared (η^2). Eta squared offers an estimate of the strength of association by utilizing only two values, the sum of squares for the treatment effect and the sum of squares for the total effect. Symbolically, eta squared is:

$$\eta^2 = \frac{SS_{\text{effect}}}{SS_{\text{total}}} \quad \text{Equation 9}$$

where SS_{effect} = the sum of squares for the treatment effect and

SS_{total} = the sum of squares for the total variance.

While eta squared is usually used to measure the strength of association between the dependent variable and the independent variable, it can be modified to measure the strength of an interaction between the independent variable and the covariate, a combination that is the basis of the homogeneity of regression assumption. Modification of the eta squared formula to the following formula allows for this measurement:

$$\eta^2 = \frac{SS_{\text{IVxCov}}}{SS_{\text{total}}} \quad \text{Equation 10}$$

where $SS_{IV \times Cov}$ = the sum of squares for the covariate.

Eta squared was chosen as the measure of the independent variable-covariate relationship because ANCOVA relies on the independence of the covariate to form the experimental treatments (Keppel, 1991). Since eta squared measures treatment effect, it can be easily applied to test the homogeneity of regression assumption. With a modified eta squared, if a large effect is detected, it can be concluded that the assumption has been violated and the results achieved from the design should be regarded with some caution.

Sample Size

Sample size is also an important factor to be considered because of its profound effect on both the outcome of ANCOVA and power. It is well known that the power of an experiment has a positive relationship with its sample size (Keppel, 1991). Maxwell, Delaney and Dill (1984) acknowledged the warning that insufficient sample size will lead to a decrease in power.

Part of the calculations for ANCOVA require the knowledge of the number of participants used within each treatment group and within the entire experiment. This necessity of sample size is seen in the degrees of freedom

for error variance. When ANCOVA is used, there is a loss of 1 degree of freedom to the error term "due to the estimation of the population slope in the calculation of the adjusted within-groups sum of squares" (Keppel, p. 312). The loss of 1 degree of freedom would not be noticed in a study using a large sample size ($n \geq 100$) but in a study using a much smaller sample ($n = 20$), the loss could mean the difference between a tested hypothesis being retained or rejected.

There is a word of caution when trying to utilize sample size to increase power. Power in ANCOVA does not rely solely on sample size but uses several factors. Rogers and Hopkins (1988a & b) give six (6) ways, including increased sample size, as a means of increasing power. There is a point that results obtained from an inflated analysis will lose their meaning. If a researcher decides, for example, to test her hypothesis using an extremely large sample (e.g. $n = 10,000$), she may find something small. Unless the researcher is looking for a precise point estimate (in which a very large sample size or a small effect size is appropriate), her results may be trivial. Further, an extremely large sample can inflate

the effect of the independent variable such that a statistically significant result is found that may otherwise be trivial.

Significance Level

Another factor that can affect the power of an experiment is its significance level. The significance level, α , is the level set prior to the experiment in which the researcher determines the dividing line between retaining the null hypothesis or rejecting it (Keppel, 1991). The convention within psychological research is to set $\alpha = .05$, since it represents the idea that a researcher is willing to reject a true null hypothesis 5 times out of 100. Significance level affects power such that if the alpha level is too stringent ($\alpha = .01$ instead of $\alpha = .05$), both the possibility of detecting a difference with the model is lessened and the probability of making a Type II error is increased. These two items, direct consequences of significance level, can lower the power of the experiment. With lack of power, whatever results are achieved can become less reliable and perhaps even detrimental to the body of research that already exists. Past research indicates that a less stringent alpha level

will help increase power (Rogers & Hopkins, 1988a; Rogers & Hopkins, 1988b). However, what is the ideal alpha level in order to maximize power?

Measurement Error

Much of the prior research examining power has been more focused on the relationship of measurement error and power. It has been shown that reducing measurement error increases power.

Rogers and Hopkins (1988a) examined the effects of measurement error and a covariate on estimates of power. Specifically, they examined how power is affected when the reliability of the dependent variable and the covariate is changed. They listed several formulas and illustrated their effectiveness by providing an example. They found six options available to a researcher for increasing power: (1) increasing the potency of treatment (effect sizes); (2) relaxing α , (3) employing a directional hypothesis; (4) increasing the sample size n ; (5) employing a more powerful statistical model; and (6) increasing the reliability of measurement of the dependent variable and the covariate (Rogers & Hopkins, 1988a). Rogers and Hopkins also created

a table that provides quick power estimates for the most common ANCOVA designs (1988b).

Williams, Zimmerman and Zumbo (1995) also found that

Power is a monotonically increasing function of reliability (defined as the proportion of observed variance that is true score variance), provided that the change in observed score variance is due exclusively to change in error score variance, whereas the function is monotonically decreasing when the standard error of measure is invariant (p. 367).

To further illustrate, Kopriva and Shaw (1991) found that the reliability of the measurement instrument can have a substantial effect on power, especially with small sample sizes were used ($n \leq 100$). Although this relationship is important and some of the research has been presented, measurement error and its effects on power will not be addressed due to the population design of the present study.

Dependent Variable-Covariate Correlations

Yet another factor that influences power in ANCOVA is the dependent variable-covariate correlation, ρ . It is well known in statistical theory that the strength of the dependent variable-covariate correlation improves the

amount of variance accounted for by the treatment effect from the total variance.

Prior research has found that increasing the correlation would lead to increased power. Rogers and Hopkins (1988a) indicated that increasing the accuracy of measurement of the dependent variable and the covariate will increase power. Kopriva and Shaw (1991) also found results that agree, noting that power increased with a high correlation.

Many researchers have used the guide that if a correlation of $r \geq .20$ is obtained, ANCOVA is the best choice despite the loss of degrees of freedom. However, Maxwell, Delaney and Dill, (1984) argued that ρ was largely irrelevant in choosing a design to increase power. They suggested that ρ should only be considered when deciding if using a covariate is worth the loss of degrees of freedom.

The issue of correlation size and its relation to power needs to be more thoroughly addressed. How high does a correlation between the dependent variable-covariate have to be in order for a marked increase is seen in power? Can a moderate correlation ($r = .50$) suffice in the decision to use ANCOVA or is a high correlation necessary ($r = .80$)?

Homogeneity of Regression

As mentioned earlier, ANCOVA has five (5) assumptions that should be met in order to have the maximum output from the design. The assumption of homogeneity of regression is that all the slopes within the cells are the same. A violation of this assumption indicates that one or more of the slopes deviates from the rest of the cells. This deviation is due to an interaction between an independent variable and the covariate. Such an interaction would lead many researchers to either transform the data or use another statistical model (Tabachnick & Fidell, 1991). However, the effect of homogeneity of regression on power has not been addressed by any of the prior research. This idea needs to be addressed. Specifically, what effect does a violation of this assumption have on power? Would a mild violation, where the slopes of a few cells deviate from the rest, have the same effect on power as a major violation, where almost all of the slopes are different? Would there be any change to power or would power remain unaffected?

There is a need to clarify the issues surrounding ANCOVA and power. Specifically, how does the degree of the dependent variable-covariate correlation (small: $r = .20$, moderate: $r = .50$ or large: $r = .80$) effect power? And how

would a change in the correlation, effect size and assumption of ANCOVA influence power?

Hypotheses

Thus, it is the intent of this study to examine the statistical power within the context of ANCOVA. It is hypothesized that as the correlation between the dependent variable and the covariate becomes larger, either in a positive or negative direction, the power of the design will increase. It is also hypothesized that the assumption of homogeneity of regression does have an effect on power such that as the heterogeneity between the slopes increases, the power will decrease. Finally, it will again be shown that as sample size for each cell increases, power will increase and that less stringent alpha levels will be associated with greater power.

METHOD

A population study was used to examine the hypotheses. Because power is based on a population parameter and not a sample statistic, analysis using a Monte Carlo study is unnecessary. By using a population study, power is calculated given specific population parameters.

Procedure

ANCOVA Design

In this thesis a one way ANCOVA model with three levels is employed. Contrast coding is assumed yielding two uncorrelated predictors. Equal n within treatment conditions is also assumed. Finally it is assumed that the covariate is independent of the independent variable (IV) and that the assumption of homogeneity of variance has been met. The equations for the ANCOVA models are presented in Appendix B.

All effect sizes are determined through calculation of appropriate eta squared estimates. Additionally, when relevant, violations of homogeneity of regression are incorporated through calculation of effect sizes for the interaction between the independent variable vectors and the covariate.

Calculating the Non-Centrality Parameter

Power is directly related to a specific null hypothesis. When this hypothesis is false the resulting test statistic is distributed as a non-central distribution that is defined by a non-centrality parameter and degrees of freedom. Larger non-centrality parameters are associated with higher estimated power.

Models that Assume Homogeneity of Regression

A Structural Equation Modeling (SEM) approach allows direct computation of the estimated population parameter relevant to power calculation: the non-centrality parameter. Using an SEM approach, a saturated model is estimated (a saturated model fits the data perfectly yielding a chi-square and degrees of freedom equal to zero). In the conditions that have met the assumption of homogeneity of regression assumption the saturated model predicts the dependent variable from the two independent variable vectors and the covariate. This saturated model can be thought of as the model associated with the alternative hypothesis (H_1).

A second model, the null model (H_0), is also estimated. In this model we assume that the effect of the independent variable vectors on the dependent variable is zero.

Therefore, this second model predicts the dependent variable from only the covariate. The null model, H_0 is a subset of the saturated model H_1 , e.g.; H_0 is nested within H_1 . Given that these are nested models, a chi-square difference test can be calculated. The resulting chi-square is the non-centrality parameter (Ullman, 1997). Power is then calculated based on this parameter.

Models that Examine the Effects of Violation of Homogeneity of Regression

Using an SEM approach a saturated model is estimated. In the conditions that examine the effects of violations of the assumption of homogeneity of regression, in the saturated model the dependent variable is predicted from the two independent variable vectors, the two IV x Covariate interactions vectors and the covariate. This saturated model can be thought of as the model associated with the alternative hypothesis (H_1).

A second model, the "wrong" null model is also estimated (H_{0-mis}). In this model we assume that the effect of the IV x Covariate vectors on the dependent variable is zero. Therefore, this second model is a misspecified model that predicts the dependent variable from only the covariate and IV vectors. This misspecified null model,

H_{0-mis} , is a subset of the saturated model H_1 , e.g., H_{0-mis} is nested within H_1 .

Under conditions of violation of homogeneity of regression, a third model is also estimated. In this model, H_0 , the DV is predicted from only the covariate. Therefore, we assume that the effect of the IV vectors on the DV is zero. H_1 is nested within H_1 and the chi square difference test between these models represents the non-centrality parameter for the test of effect of the IV vectors on the DV given violation of the assumption of homogeneity of regression. Power is then calculated based on this parameter.

Variables

There were four independent variables and one dependent variable. The dependent variable was the calculated power. The first independent variable was the correlation between the dependent variable and the covariate in the design. This was divided into 11 levels ($r = \pm 1.00, \pm 0.80, \pm 0.50, \pm 0.30, \pm 0.10, 0.00$) in order to examine the hypothesis that as the dependent variable-covariate correlation increases, either in a positive direction or a negative direction, power will increase.

The second independent variable was the assumption of homogeneity of regression which was divided into three levels (assumption was met, assumption with a mild violation and assumption with a gross violation) in order to examine the hypothesis that the greater the heterogeneity between the cells, the less power in the design. Values for the three levels were determined by effect sizes for the independent variable-covariate interaction (η^2). The first level was $\eta^2 \leq .05$; the second level was $\eta^2 = .30$; the third level was $\eta^2 = .50$.

The third independent variable was sample size for each cell, which was divided into five levels ($n = 10, 20, 30, 40, \text{ and } 50$) in order to show that a larger n will increase power.

The fourth independent variable was significance level, which was divided into two levels ($\alpha = .05$ and $\alpha = .01$ in order to show that the less stringent the alpha level, the greater the power.

Design

The design of interest was a one-way ANCOVA with three levels. Treatment effects were set at $\eta^2 = .20$ and $\eta^2 = .10$. Two different sets of effects were used for the matrices

(See Appendix B) in EQS because the program would not calculate the non-centrality parameter the same effect size for both vectors.

Initial tests of the hypotheses were run with treatment effects set at $\eta^2=.20$ and $\eta^2=.10$, respectfully. However, it was discovered that the power calculations yielded from these treatment effects would be too high to demonstrate any differences between the hypotheses. It was therefore decided that the set treatment effects should be reduced to $\eta^2=.10$ and $\eta^2=.05$.

Further, to gain a better picture of the effects on power, the number of correlations was expanded to include $r = \pm 0.20, \pm 0.40, \pm 0.60, \pm 0.70$, and ± 0.90 . This changed the design from an $11 \times 3 \times 5 \times 2$ matrix to a $21 \times 3 \times 5 \times 2$. Finally, to further demonstrate the hypothesized effect of homogeneity of regression on power, two additional levels of the independent variable-covariate interaction were added ($\eta^2=.40$ and $\eta^2=.70$). Thus the matrix was changed from a $21 \times 3 \times 5 \times 2$ to a $21 \times 5 \times 5 \times 2$ matrix.

During the power calculations, it was found that the resulting power estimates were identical for both the positive and negative correlations. Based on this

observation, it was decided that only the positive correlations would be used during the actual power calculations and the results would be generalized to the negative correlations. It was also found that some combinations of the variables yielded a negative determinant. These combinations were included in the data tables but were excluded in the graphs.

RESULTS AND DISCUSSION

Appendix C presents the results from the hypotheses concerning sample size, dependent variable-covariate correlations and significance level. For each figure, power values are labeled on the y-axis and the correlations (r) are labeled on the x-axis. Only positive values for the correlations are displayed though the power value results apply to both positive and negative correlations. The legend contains the lines coded by sample size with the total sample size in parentheses. Power values are graphed according to sample sizes. Numeric power values are displayed immediately below the correlation values.

Sample Size

The hypothesis that increasing the sample size for each cell of an ANCOVA design will increase power was confirmed. In all the figures presented in Appendix C, power values increased when the sample size was increased. By examining Figures 1-5, where the alpha level was held at .01 for all trials, power increased when n increased from 10 per cell to 50 per cell. Figure 1 clearly illustrates this trend. In this figure, the starting points for the power values were different based on sample size. When $n = 10$, the power was .248924 versus when $n = 30$, the power was

.814331 and a further increase was seen when $n = 50$; power at this sample size started at .976433. Figures 2,3,4, and 5, all with the same level of alpha, also show the same pattern of results.

When alpha level was set at .05, Figures 6,7,8,9, and 10 showed the same type of increase occurred. Figure 6 shows that when $n = 10$, power was low (.473939) at $r = 0.00$ but increased when $n = 50$ (.995144) for the same r .

These results agree with prior research that increasing sample size is an effective way to increase power (Keppel, 1991; Kopriva & Shaw, 1991; Cohen, 1988; Rogers & Hopkins, 1988a; Rogers & Hopkins, 1988b). However, there is a limit to how much sample size can be increased to gain power. Since power is "the probability that a significant effect will be found in [a] proposed study" (Rogers & Hopkins, 1988b), the more power in a design, the more likely the researcher is to detect a difference if a true difference exists. Although Figure 1 illustrates how power increases, the amount of that increase is not as substantial at $n = 50$ than when $n = 30$. It is shown that both sample sizes increase, however the potential for power is at its maximum with $n = 50$. For

$n = 30$, the increase is much more noticeable and therefore more reachable. Absolute power is a hypothetical ideal, one that is seldom achieved. Thus, the psychological field has a general consensus that moderate power ($1 - \beta = 0.70$) is acceptable though high power ($1 - \beta > 0.80$) is preferred. Given this guideline, it is recommended that with all other factors held constant, sample size for each cell be set at 30. If all other factors cannot be held constant, a slightly higher cell sample size of 40 is recommended. With $n = 40$ per cell, even if other factors reduce power (such as a low dependent variable-covariate correlation) the cell sample size will allow for a better statistical chance of detecting a true difference if one exists.

Significance Level

The hypothesis that using a more stringent significance level will decrease power was also confirmed by the present study. Two common alpha levels were utilized in this study and the results can best be seen in comparing the power values from Figures 1 and 6. Figure 1 has power starting at .248924 where Figure 6 has power starting at .473939. Although Figure 1, with an alpha level set at .01, shows the greatest increases of power, it

does not negate the point that power is low, lower than the same power value under the same conditions for $\alpha = .05$.

The changes in power can also be seen when comparing Figures 3 and 8. When the alpha level is set at .05, power approaches its maximum value (1.000000) faster over more sample sizes (three sample sizes for $\alpha = .05$ versus two sample sizes for $\alpha = .01$) than when the alpha level is set at .01. As further evidence, when $n = 50$ at $\alpha = .05$ the values above .990000 are achieved when $r = 0.00$ whereas for the same conditions at $\alpha = .01$, values above .990000 are achieved when $r = \pm 0.40$.

The clearest illustration of the influence of significance level on power can be seen when Figures 5 and 10 are compared. In Figure 5, power values start at .253136 but in Figure 10 power values start at .479159. All calculated power values start higher in Figure 10, with an alpha level set at .05, than in Figure 5, with alpha level set at .01. Both figures show that the power values increase at the same rate, but that does not negate that power values for the significance level of .01 are lower than for the same conditions at .05.

These results agree with Rogers and Hopkins (1988a; 1988b) who cited that relaxing the alpha level, for example from .01 to .05, would increase power. The logic behind this fact lies in the meaning of the alpha level itself. Since alpha level is the researcher's guide to the number of trials that the null hypothesis would hold true versus the number of trials that it would be rejected, a less stringent alpha level allows more acceptance of the number of trials that a researcher is willing to be wrong. This acceptance leads to increased power because there is more flexibility to the experiment even though there is an increased opportunity for the obtained results to be false.

The present research, however, does not imply that power will not increase with a more stringent alpha level. Rather, power will increase but the rate of increase is not as strong at $\alpha = .01$ as it would be at $\alpha = .05$. This relationship can be seen when Figures 1 and 6 are compared. In this observation, Figure 1 has power values that start low but end high as the correlations approach ± 1.00 . Figure 6 also shows the same increase in power values but the initial value is higher and the approach to absolute power is quicker than the same conditions with $\alpha = .01$.

These differences in power values are consistent throughout the remaining figures. Figures 2 and 7, when compared, have initial power values that start from .253136 and .479159, respectively. This same difference continues through the comparisons of Figures 3 and 8, Figure 4 and 9, and Figures 5 and 10. Each of these pairs of figures contains a lower starting power value for $\alpha = .01$ than for $\alpha = .05$.

Given these patterns of results and prior research, it is concluded that the alpha level for an ANCOVA design be set at .05 in order to aid the applied researcher in reaching the goal of acceptable power, provided that the increase in Type I error is worth the risk.

Dependent Variable-Covariate Correlations

The hypothesis that increasing the correlation between the dependent variable and the covariate would increase the power was confirmed. All experimental trials utilizing correlations yielded increasing power values. Figure 1 illustrates one of the conditions which correlations were used. Examination of any line of data from Figure 1 indicates that power increases from a low value (.248924 for $n = 10$, $r = 0.00$) to a high value (.999895 for $n = 10$, r

± 0.90). Figure 6 also yielded the same pattern of results. For $n = 10$, the power value was low (.473939), however as the correlations increased so does the power. For both Figures 1 and 6, power values stop at $r = \pm 0.90$; no values could be obtained for $r = \pm 1.00$. The matrices used for the chi-square difference tests reached a negative determinant at $r = \pm 1.00$. A negative determinant states that there is more covariance than variance in the matrix and the calculation results in a negative number. This limit for negative determinant was reached sooner as the conditions were changed, e.g., $r = \pm 0.70$ for $\eta^2 = .30$; $r = \pm 0.60$ for $\eta^2 = .40$; $r = \pm 0.50$ for $\eta^2 = .50$ and $r = \pm 0.30$ for $\eta^2 = .70$, but in each set of conditions the power values continued to increase. In Figure 2, the power increases from .933357 at $n = 40$ to 1.000000 but reaches the limit of power at $r = \pm 0.70$ for all levels of n . In Figure 3, power starts low for every $r = 0.00$ but increases until reaching its limit at $r = \pm 0.60$. For Figures 4 and 5, power still increases but reaches its limit much sooner ($r = \pm 0.50$ and $r = \pm 0.40$, respectively). All the limits of power values stop at the same correlations for Figures 7 through 10.

The results of the present study agree with prior research done in the area (e.g. Rogers & Hopkins, 1988a; Rogers & Hopkins, 1988b). These researchers suggested that using one or more covariates that correlate with the dependent variable would increase power. Rogers & Hopkins (1988a & 1988b) also provided basis and agreement with the results of a later study done by Kopriva & Shaw (1991) and both studies rebuked Maxwell, Delaney and Dill (1984) who stated that the dependent variable-covariate correlation (ρ) only be used to decide whether or not a covariate was worth the cost of degrees of freedom.

Thus it is concluded in the present study that to achieve moderate power a minimum correlation of $r = \pm 0.50$ be used when the sample size is small ($n = 20$ per cell). With larger sample sizes ($n \geq 30$ per cell), a lower correlation of $r = .20$ can be used. Under conditions where there is a potential problem (e.g., low sample size or a mild violation to homogeneity of regression), a higher correlation of $r = \pm 0.70$ is recommended. Any other conditions that were used in the present study (e.g., severe heterogeneity of regression) that occur in a

psychological experiment, it is recommended that ANCOVA not be used but some other statistical design be considered.

Homogeneity of Regression

Appendix D presents the results from the hypothesis concerning homogeneity of regression. Power values are labeled on the y-axis and the dependent variable-covariate correlations (r) are labeled on the x-axis. Only positive values for the correlations are displayed though the power value results apply to both positive and negative correlations. The legend contains the lines coded by violation to the homogeneity of regression assumption (e.g., E0.05 for $\eta^2 \leq .05$; E0.30 for $\eta^2 = .30$). Power values are graphed according to their results from the homogeneity of regression violation. Numeric power values are displayed immediately below the correlation values.

The hypothesis that violating the assumption of homogeneity of regression for ANCOVA would lead to a decrease in power was not supported in the present study. Five levels of the violation to homogeneity of regression were used in the present study - no violation ($\eta^2 \leq .05$), mild violation ($\eta^2 = .30$ and $\eta^2 = .40$), moderate violation ($\eta^2 = .50$) and severe violation ($\eta^2 = .70$). Figures 11

through 20 show that as the level of violation increased, there was no change to the power values. Each level of the independent variable-covariate interaction had no effect on power, no matter what conditions were implemented. Figures 11 and 16 illustrate that the lines of the graph overlap and the numeric power values in the tables display only trivial differences. Even when the sample sizes were increased with the increase in the violation, no change was observed in the power values. Figures 11, 12, and 13 also show that the power values stopped at the same points for the correlations as in Figures 1-3 for the same reason—a negative determinant was reached in the matrices. Figure 11 shows how power remained unaffected, starting low for $r = 0.00$ (.248924 for $\eta^2 \leq .05$ and .253136 for all other levels of η^2) and ending high. This same pattern can be seen for Figures 16, 17, and 18, where the alpha level was set at .05.

As noted under the sections concerning sample size and significance level, power continues to increase with no hindrance from any violation to the homogeneity of regression assumption. This can be seen in Figures 15 and 20, where the violation is at its most severe. Power

values are still high in Figure 20, starting at .995144 for $\eta^2 \leq .05$ and .995566 for all other levels of η^2 . This pattern of unaffected power values can be seen in Figures 12 and 17, 13 and 18, as well as 14 and 19.

This pattern of results initially seems to refute the suggestions made by Tabachnick and Fidell (1991) that another statistical design should be used when there is a violation of the homogeneity of regression assumption. However, ANCOVA is affected by the violation. This can be seen within the F ratio when it is converted in terms of eta-squared. When changed, the F ratio is as follows:

$$\frac{\eta^2_{IV,Cov} - \eta^2_{Cov} / df}{1 - \eta^2_{IV,Cov} / df} \quad \text{Equation 11}$$

The converted F ratio is still explained variance over unexplained variance, but using eta-squared demonstrates the only effect homogeneity of regression has on ANCOVA. When there is a violation to the assumption of homogeneity of regression, there is no place to partition out the independent variable-covariate interaction from the unexplained variance. Ideally, the unexplained variance has random error. With the addition of the independent variable-covariate interaction, systematic error is added

to the unexplained variance. It is this addition of systematic error, found to be the only effect homogeneity of regression has in the present study, that causes the interpretation problem of the results. With an interaction, there is no indication of what effected the dependent variable in the experiment. Was it the independent variable or was it the covariate? Which had the greater effect? Further, any results obtained under these conditions would be difficult, if not impossible, to generalize because there would be no way to guarantee that it was the independent variable that effected the dependent variable.

This is made clear when an experimental design utilizing any kind of pharmaceutical(s) is considered. In clinical trials, when the effects of a drug are the goal of the research, an interaction between the wrong variables (e.g. between the independent variable and the covariate) could lead to misinformation that can have potentially serious consequences.

Thus it is concluded by this research that a violation of the homogeneity of regression would exclude ANCOVA as a statistical possibility. The whole focus of ANCOVA is its ability to better isolate the relationship between an

independent variable and a dependent variable by using a covariate to statistically correct for some variance within the design. If a violation is suspected, it is recommended that another statistical model be utilized, one that does not rely solely on the use of a covariate to achieve more power.

Overall Conclusions for the Applied Researcher

Based on the results from the four hypotheses, a combination can be made to determine the optimum levels in order to maximize power. As can be seen from Figures 6 through 10 and Figures 16 through 20, an alpha level of .05 increases power. To further increase power, a sample size of 30 participants in each cell or 90 participants total should yield enough power from a one-way ANCOVA to be well-received within the psychological fields. If 30 per cell is impractical, a sample size of 20 per cell will be sufficient, provided there is a dependent variable-covariate correlation of at least $r = \pm 0.40$. It is also recommended that there be no violation of the homogeneity of regression assumption, not because power will be effected but to allow for clearer interpretation and replication of the results.

All of these recommendations, save the homogeneity of regression assumption, agree with prior research (Tabachnick & Fidell, 1996; Rodgers & Hopkins, 1988a, Rogers & Hopkins, 1988b). Now that homogeneity of regression has been added to the research done on ANCOVA, a new dimension has been opened. This new avenue allows the psychological community more opportunity to increase the power of the ANCOVA design. With continued advances into the perfection of statistical models, psychologists will be able to increase their understanding of the thoughts, actions and reasons of the individuals and groups that comprise the human race.

APPENDIX A: Design of Experiment

Figure 1

Model of design with sample sizes (n), $\alpha = .01$, DV-covariate correlations and IV-covariate interactions (η^2).

Figure 2

Model of design with sample sizes (n), $\alpha = .05$, DV-covariate correlations and IV-covariate interactions (η^2).

APPENDIX B: Sample EQS Input Files

Figure 1

Sample program file for testing no violation of homogeneity
of regression

```
/TITLE
ANCOVA no violation of H of regression
/SPECIFICATIONS
  VARIABLES= 4; CASES= 90
  METHOD=ML
/LABELS
V1=IV1; V2=IV2; V3=cov; V4=dv
/EQUATIONS
!Run 2nd with just DV-v4 and COV-v3;
V4 = *V3 + E4
!1st run the below equation with IV and Cov -should be zero
!v4 = *v1 + *v2 + *v3 + e4
/variances
  e4 = *
  v1, v2, v3 = *;
/Matrix
1.00      0      0      .316
0      1.00      0      .22
0      0      1.00      .90
.316      .22      .90      1.00

/end
```

Figure 2

Sample program file for testing violation of the
homogeneity of regression assumption

```
/TITLE
ANCOVA violation of H of regression r square = .30
/SPECIFICATIONS
  VARIABLES = 6;  CASES= 150;
  METHODS=ML
/LABELS
V1=IV1; V2=IV2; V3=cov; V4=dv; V5=IV1_COV; V6=IV2_COV;
/EQUATIONS
!Run 3rd with just Dv-v4 and Cov-v3
  V4 = *V3 + E4;
!2nd run the below equation with IV and Cov
  !v4 = *v1 + *v2 + *v3 + e4
!1st run with the below equation
  !V$ = *V1 + *V2 + *V3 + *V5 + *V6 + e4;
/variances
e4 = *;
v1, v2, v3 = *;
v5, v6 = *;
/Matrix
1.00      0      0      .316      0      0
0      1.00      0      .224      0      0
0      0      1.00      .80      0      0
.316      .224      .80      1.00      .44      .32
0      0      0      .44      1.00      0
0      0      0      .32      0      1.00

/end
```


APPENDIX C: Power Results by Sample Size,
Significance Level and Correlations.

Figure 1

Power values across DV-Covariate correlations (r) and
sample sizes for $\alpha = .01$ and IV-Covariate interaction
(η^2) $\leq .05$.

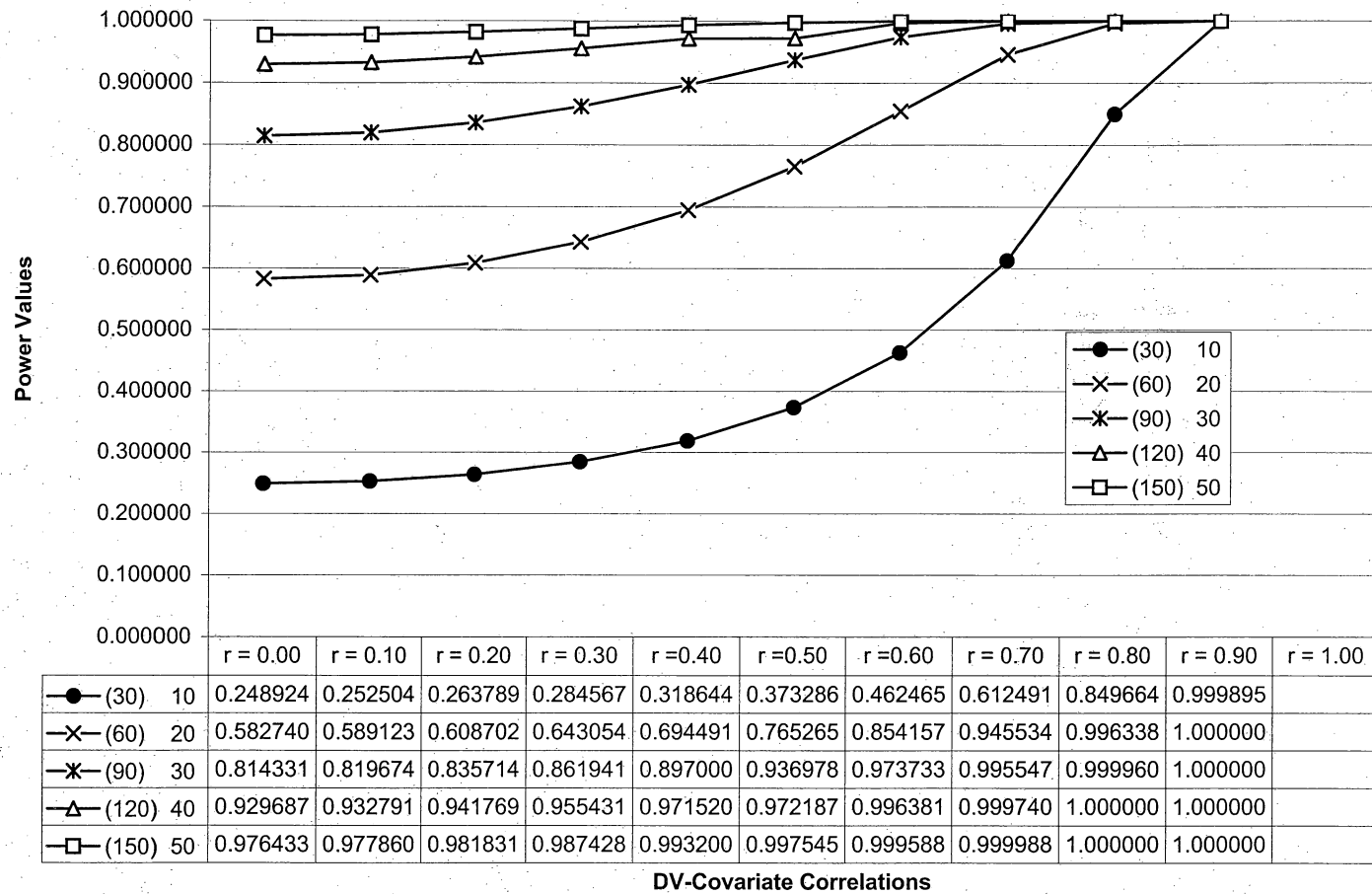


Figure 2

Power values across DV-Covariate correlations (r) and sample sizes for $\alpha = .01$ and IV-Covariate interaction (η^2) = .30.

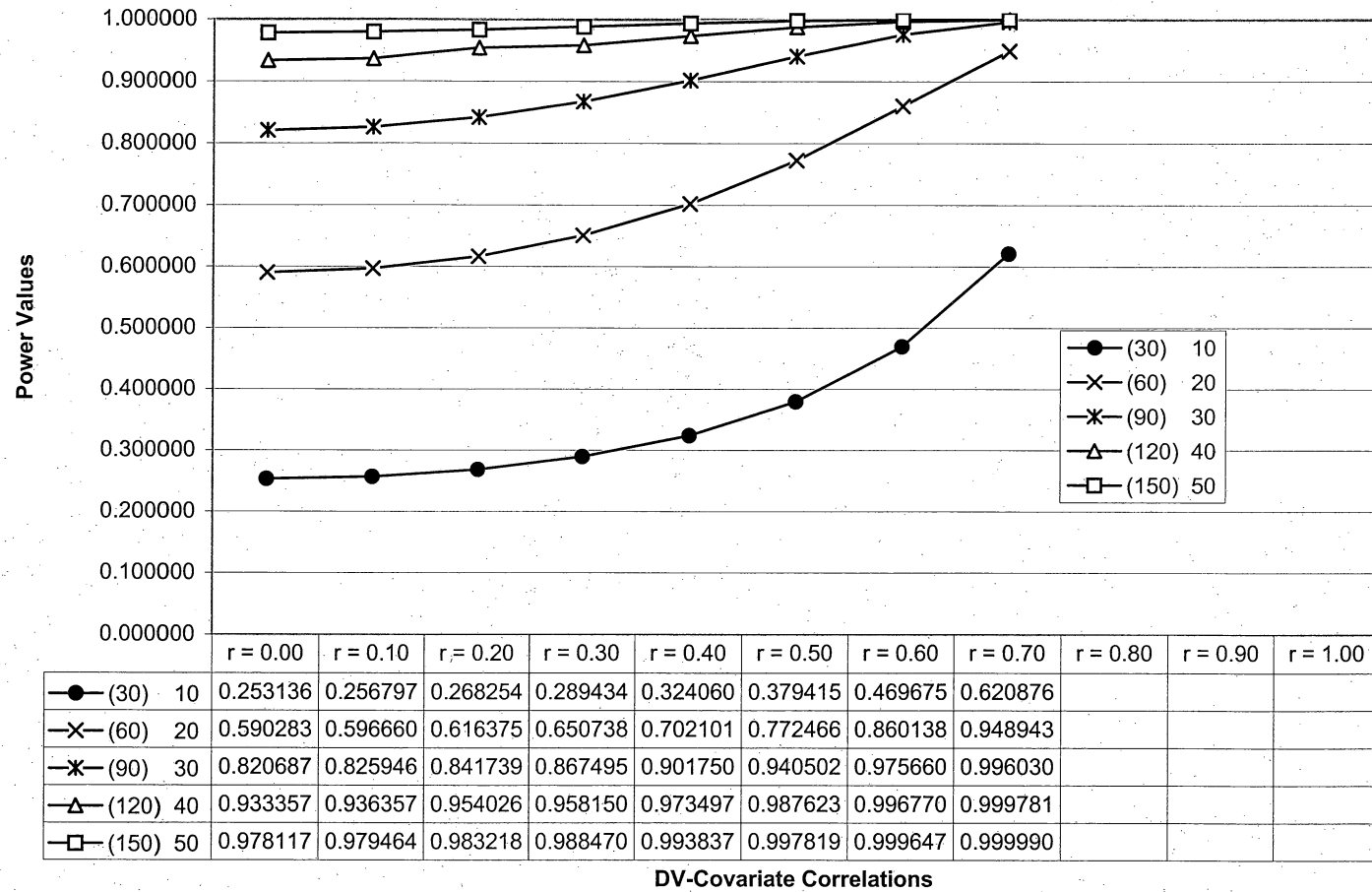


Figure 3

Power values across DV-Covariate correlations (r) and sample sizes for $\alpha = .01$ and IV-Covariate interaction (η^2) = .40.

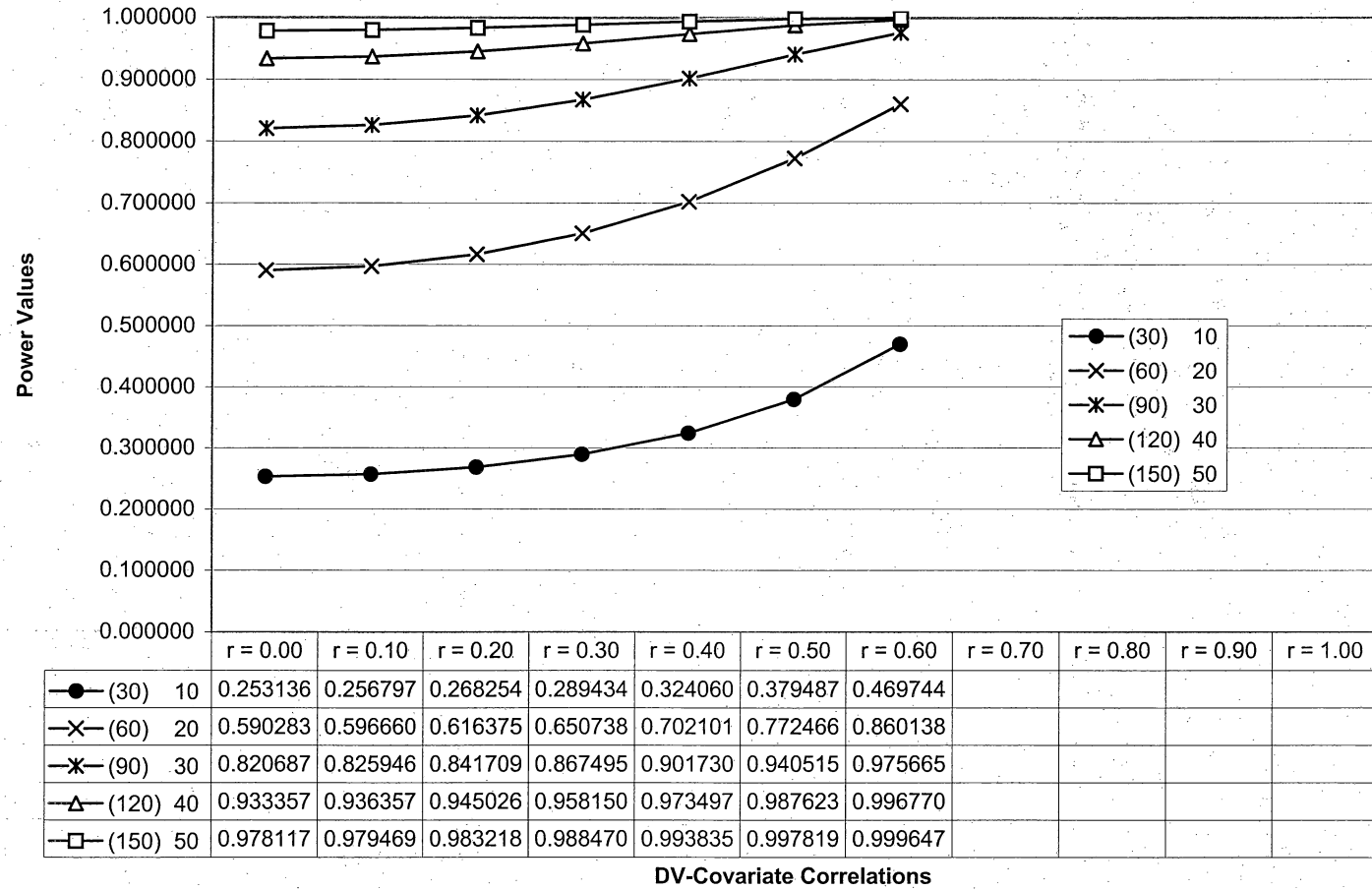


Figure 4

Power values across DV-Covariate correlations (r) and sample sizes for $\alpha = .01$ and IV-Covariate interaction (η^2) = .50.

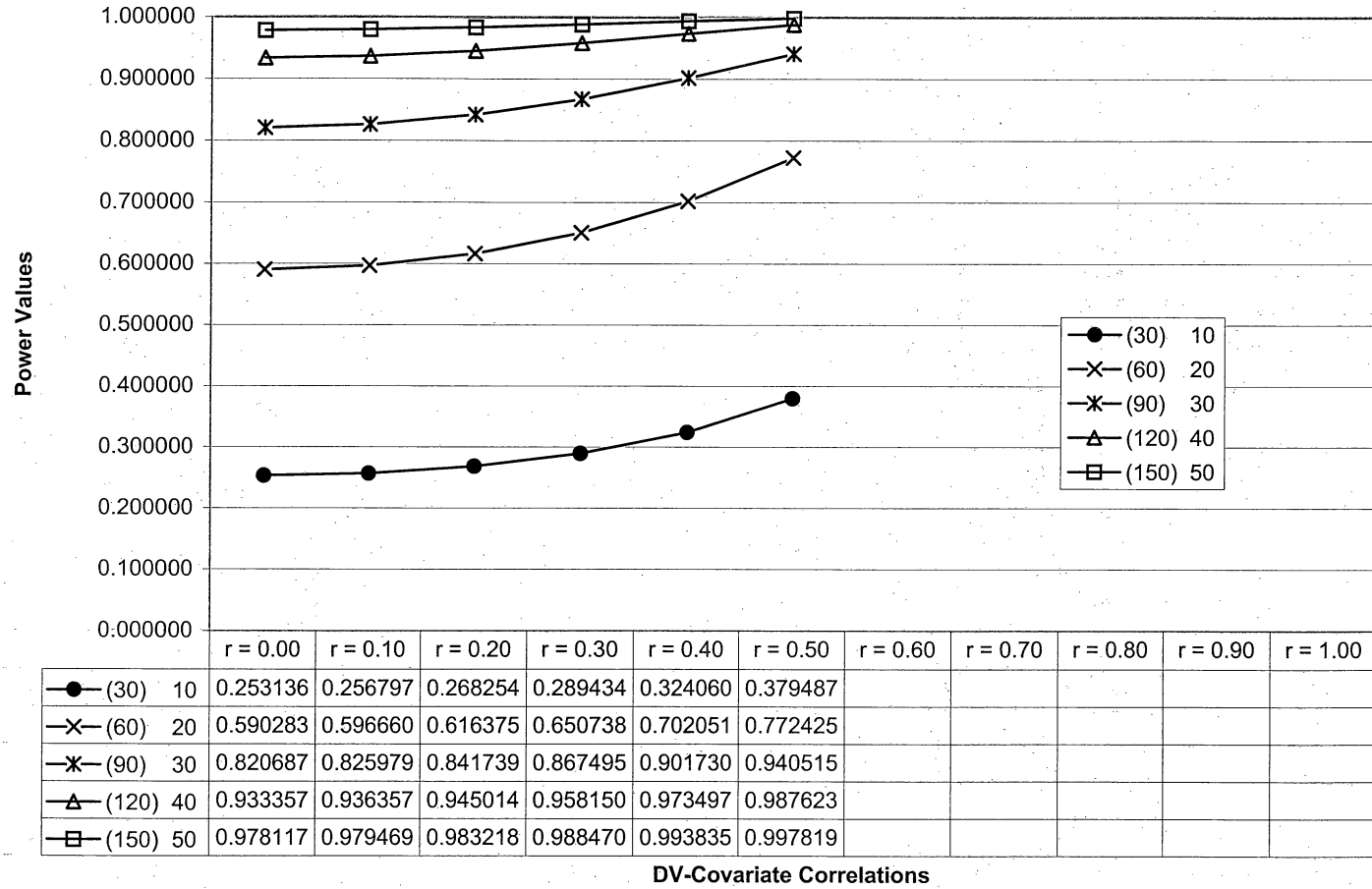


Figure 5

Power values across DV-Covariate correlations (r) and sample sizes for $\alpha = .01$ and IV-Covariate interaction (η^2) = .70.

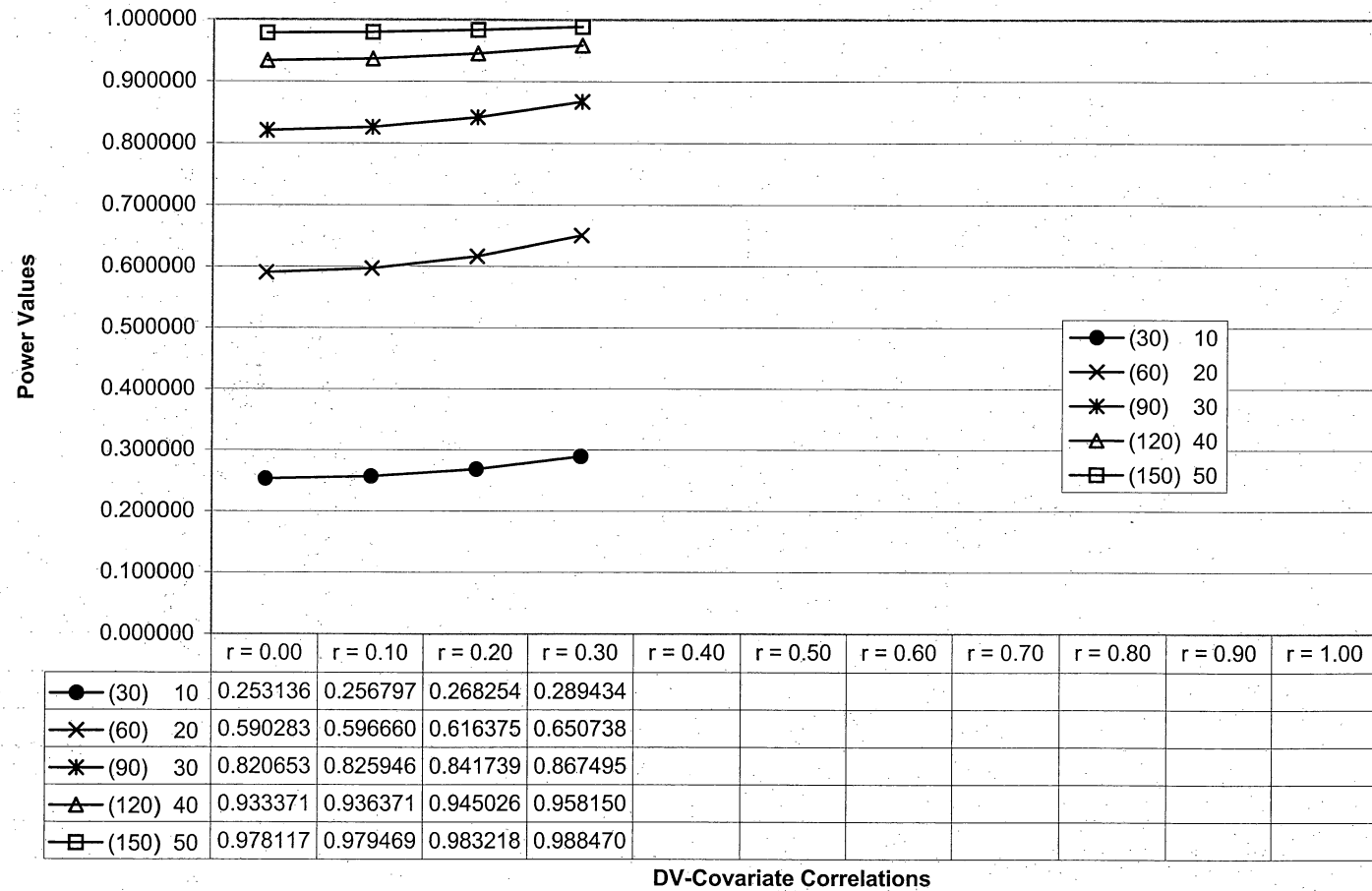


Figure 6

Power values across DV-Covariate correlations (\mathbf{r}) and sample sizes for $\alpha = .05$ and IV-Covariate interaction $(\eta^2) \leq .05$.

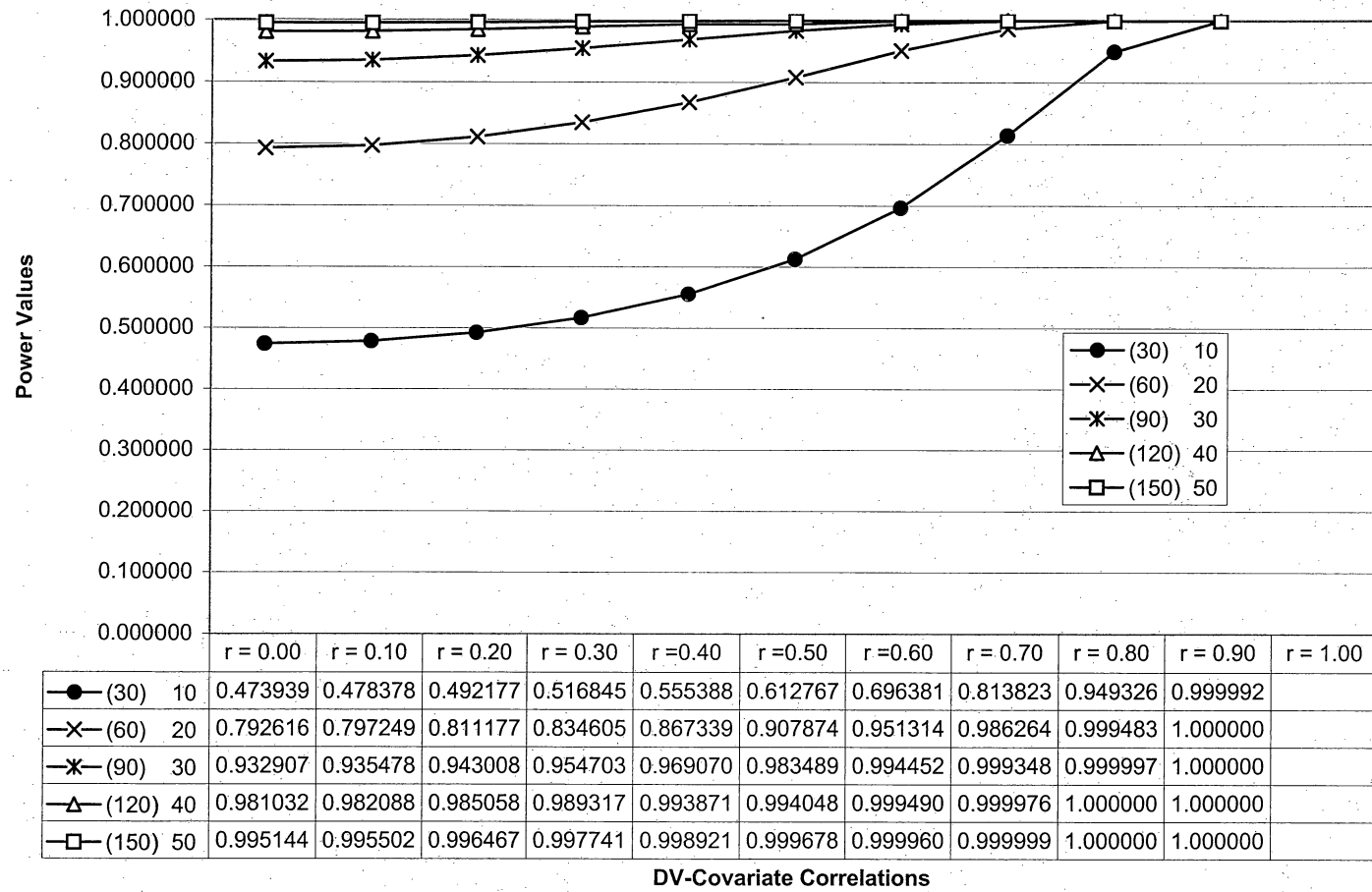


Figure 7

Power values across DV-Covariate correlations (r) and sample sizes for $\alpha = .05$ and IV-Covariate interaction (η^2) = .30.

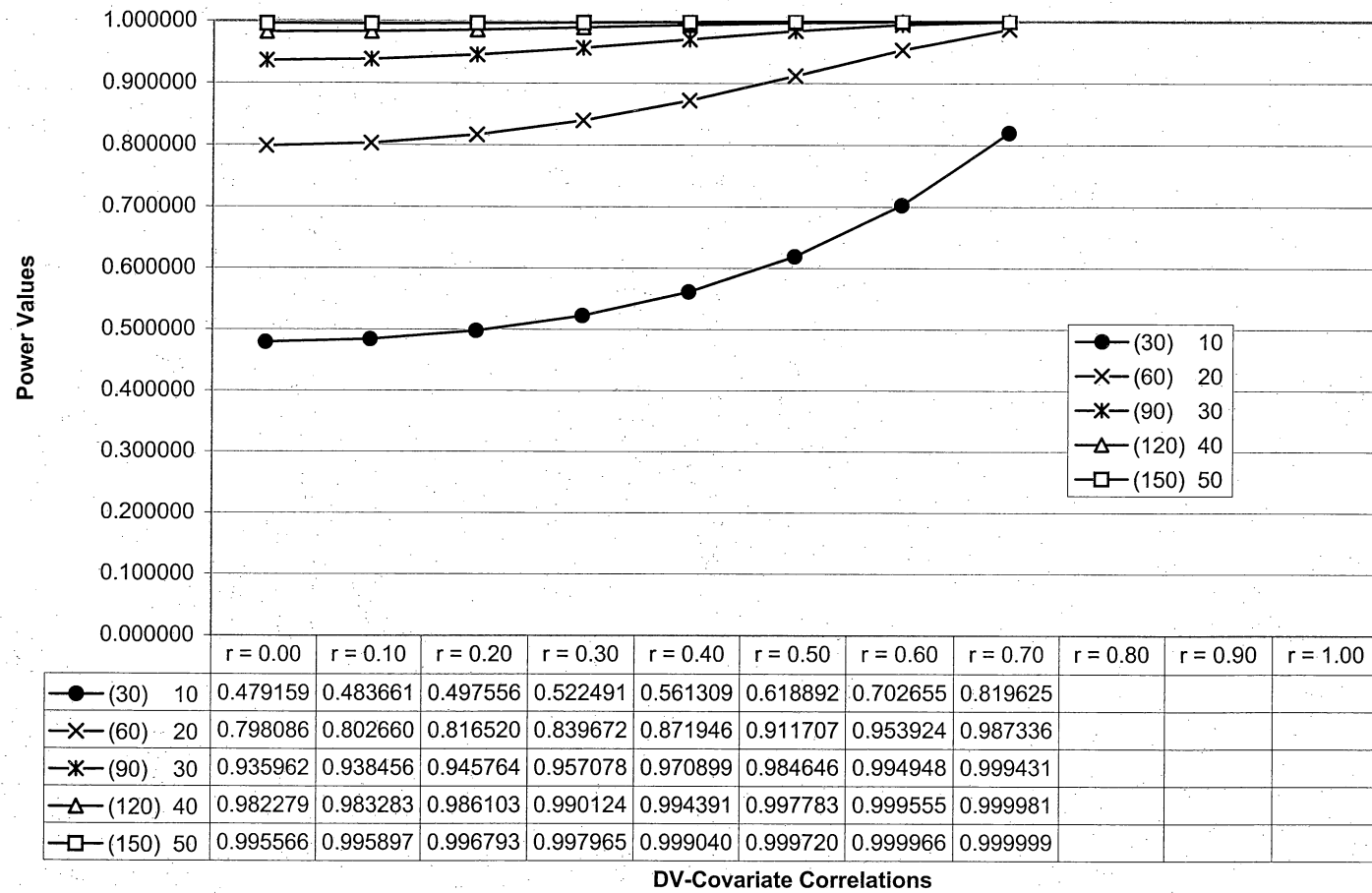


Figure 8

Power values across DV-Covariate correlations (r) and sample sizes for $\alpha = .05$ and IV-Covariate interaction (η^2) = .40.

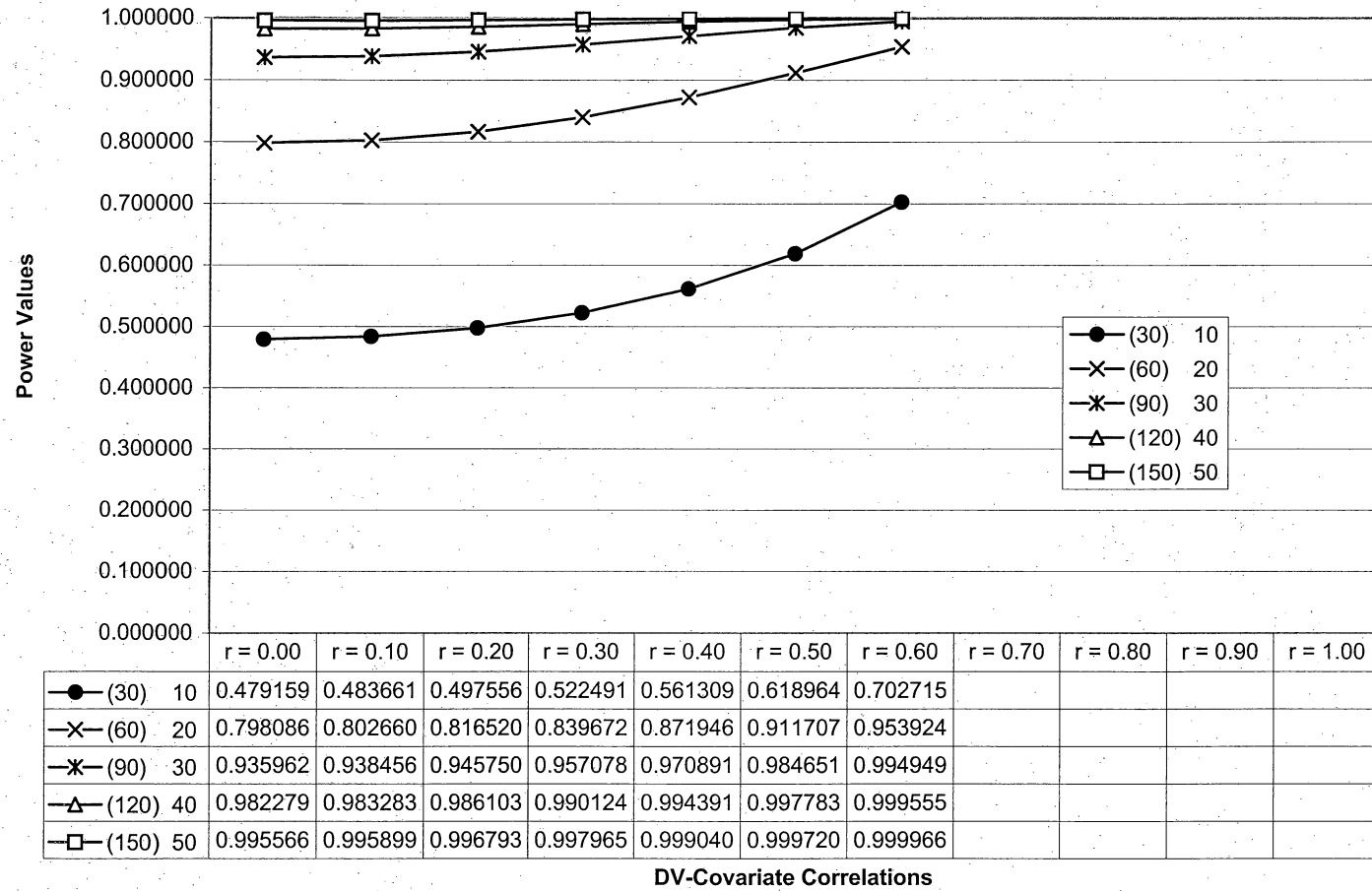


Figure 9

Power values across DV-Covariate correlations (r) and sample sizes for $\alpha = .05$ and IV-Covariate interaction (η^2) = .50.

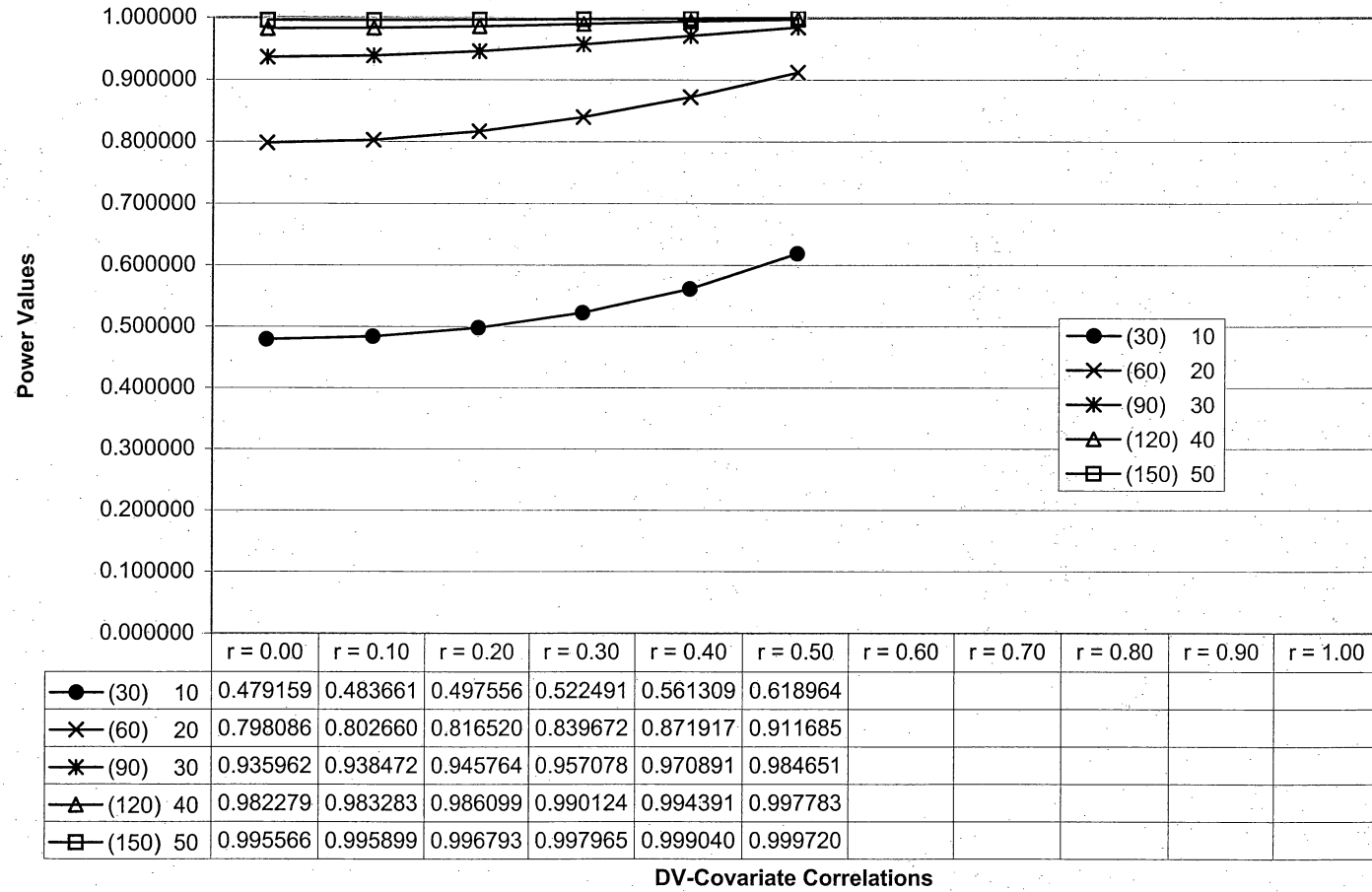
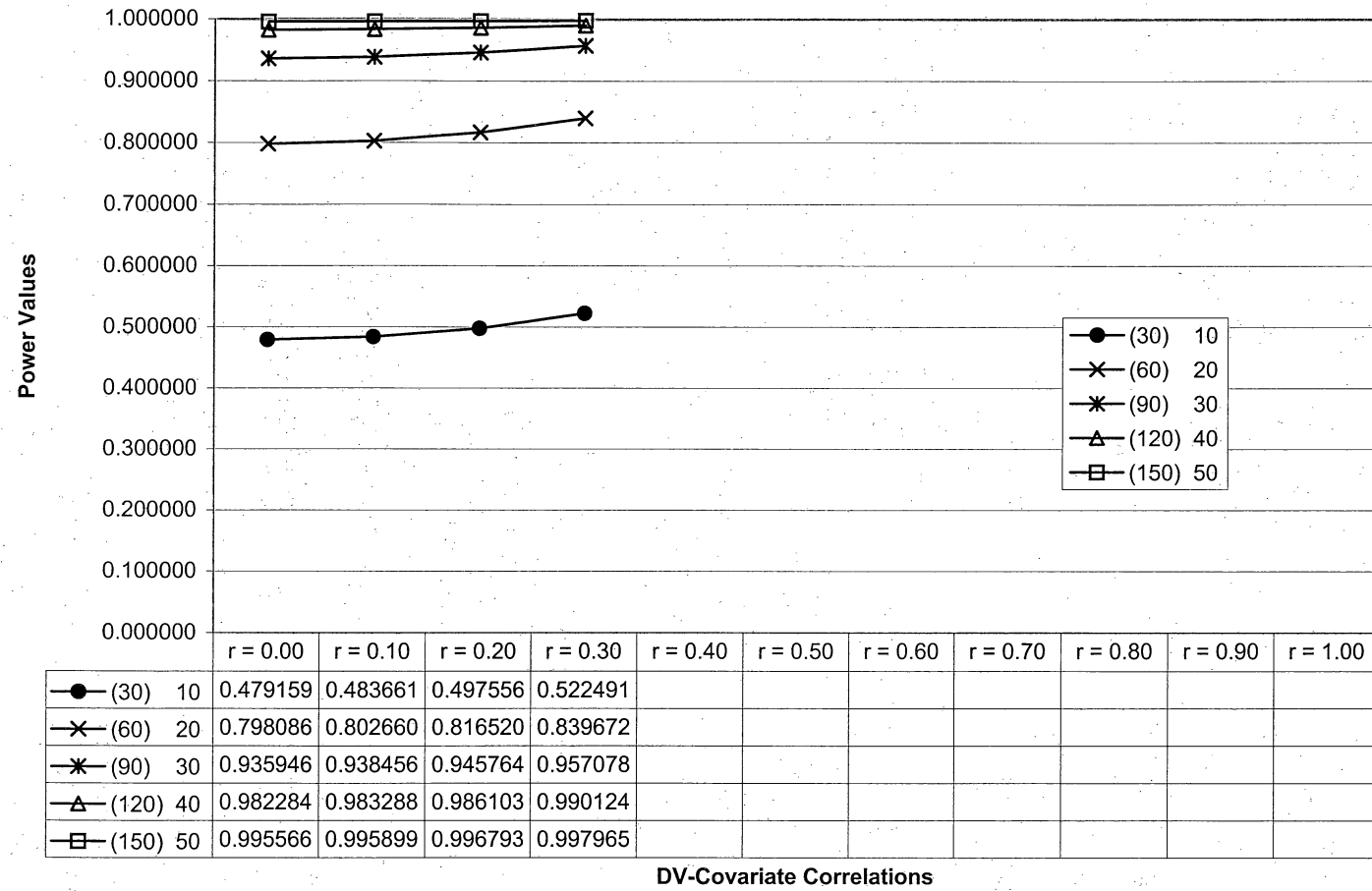


Figure 10

Power values across DV-Covariate correlations (r) and sample sizes for $\alpha = .05$ and IV-Covariate interaction (η^2) = .70.



APPENDIX D: Power Results for Homogeneity of Regression

Figure 11

Power values across DV-Covariate correlations (r) and IV-covariate interactions (η^2) for $\alpha = .01$ and $n = 30$.

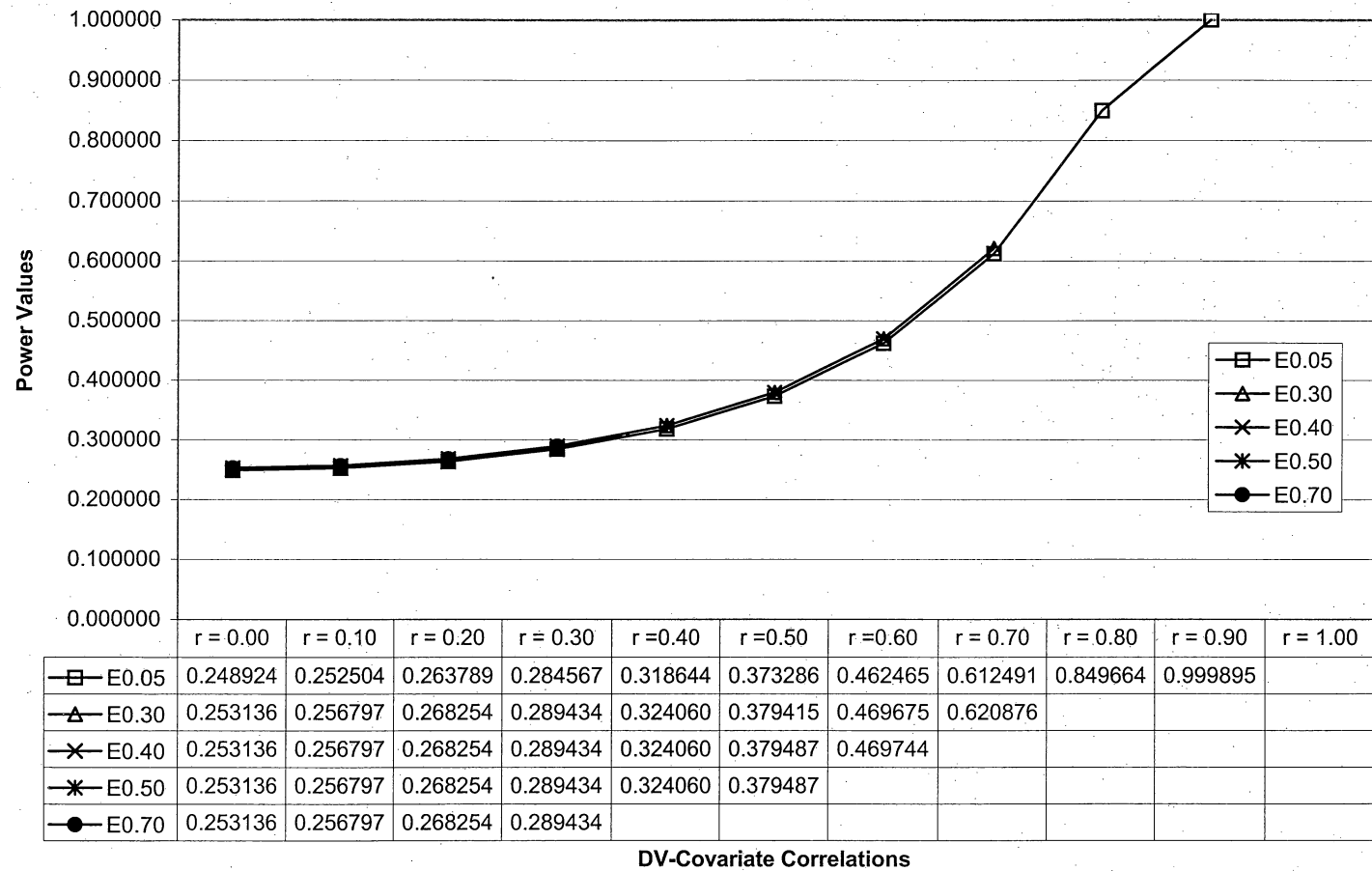


Figure 12

Power values across DV-Covariate correlations (r) and IV-covariate interactions (η^2) for $\alpha = .01$ and $n = 60$.

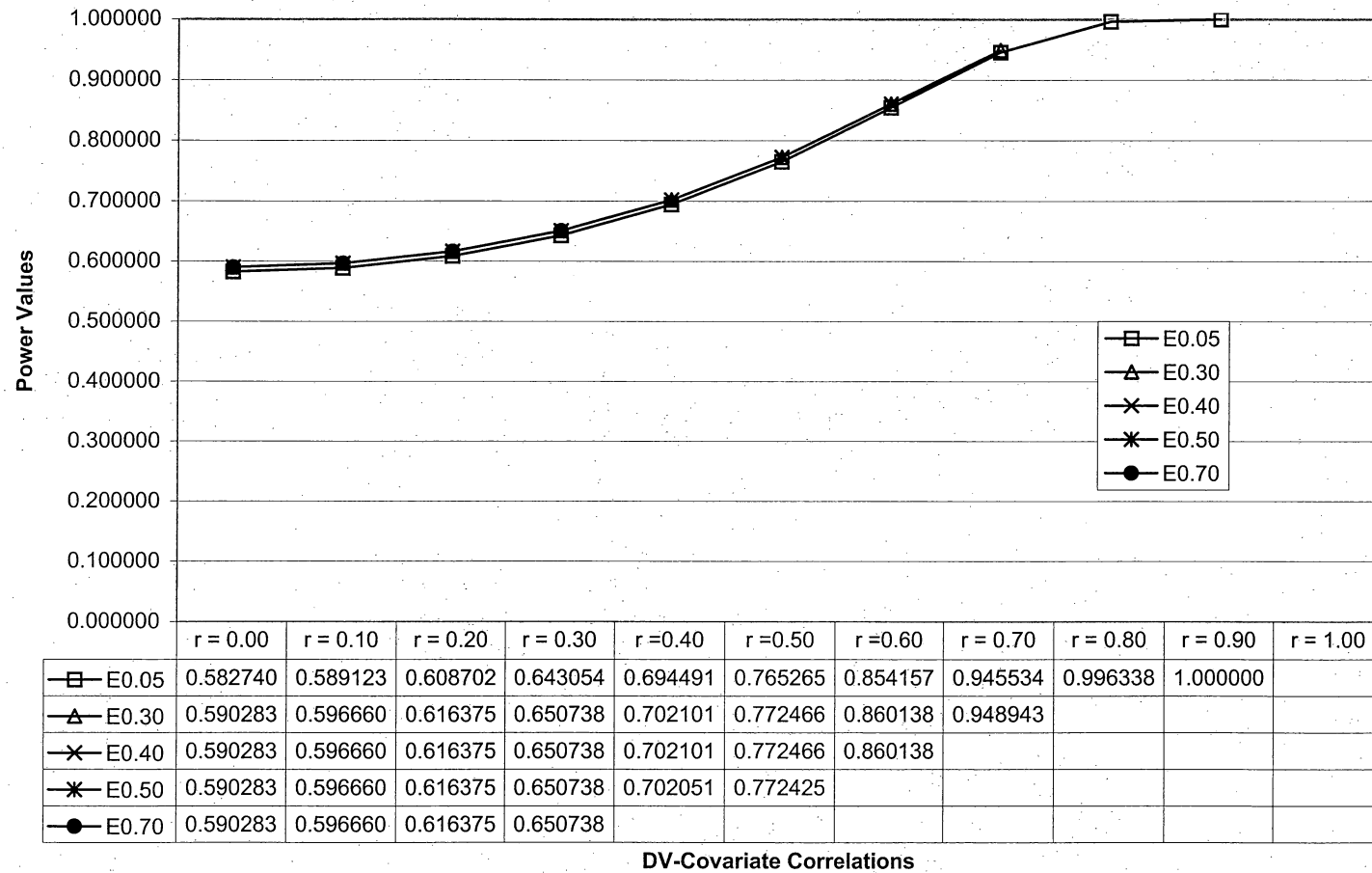


Figure 13

Power values across DV-Covariate correlations (r) and IV-covariate interactions (η^2) for $\alpha = .01$ and $n = 90$.

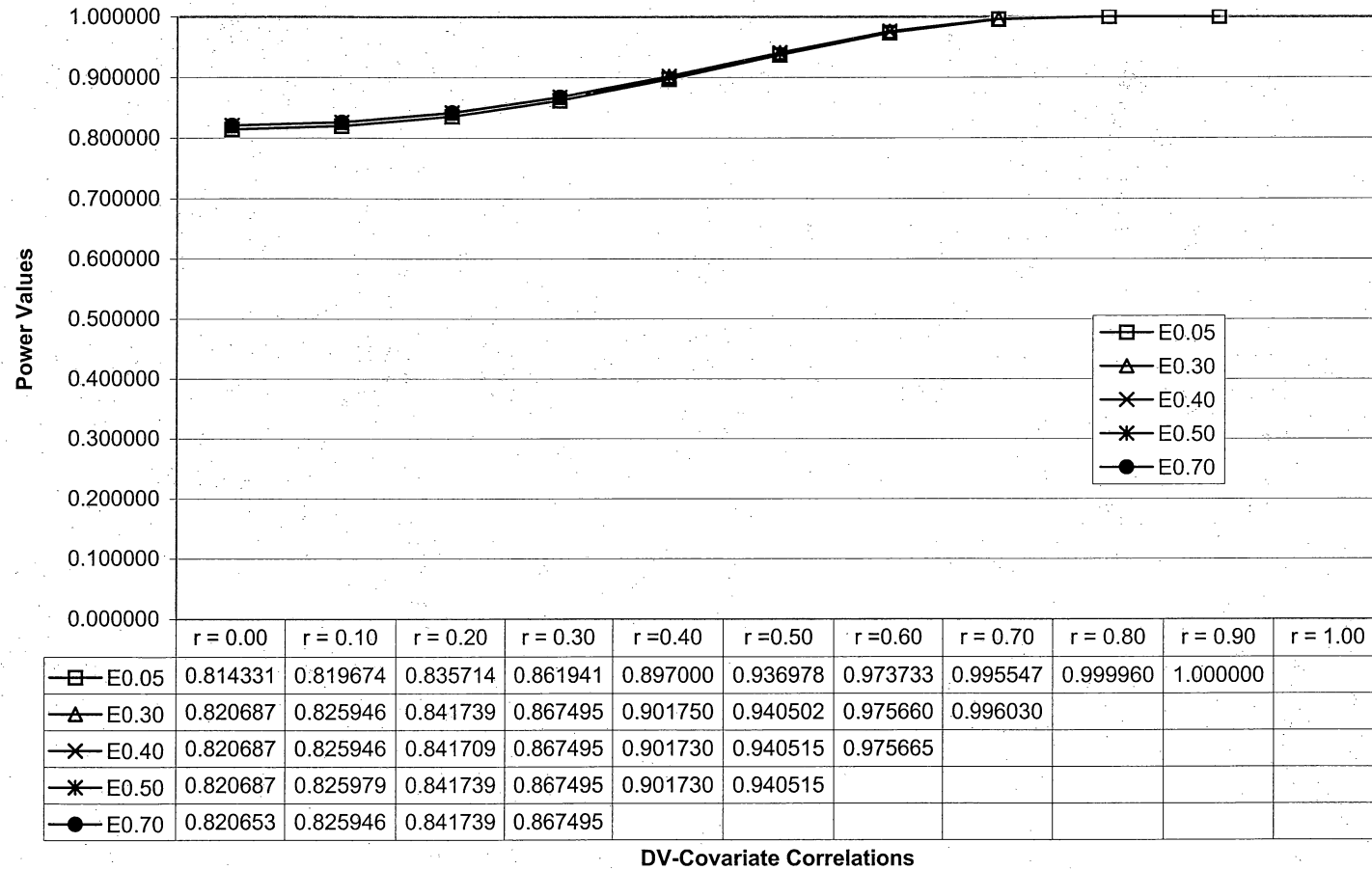


Figure 14

Power values across DV-Covariate correlations (r) and IV-covariate interactions (η^2) for $\alpha = .01$ and $n = 120$.

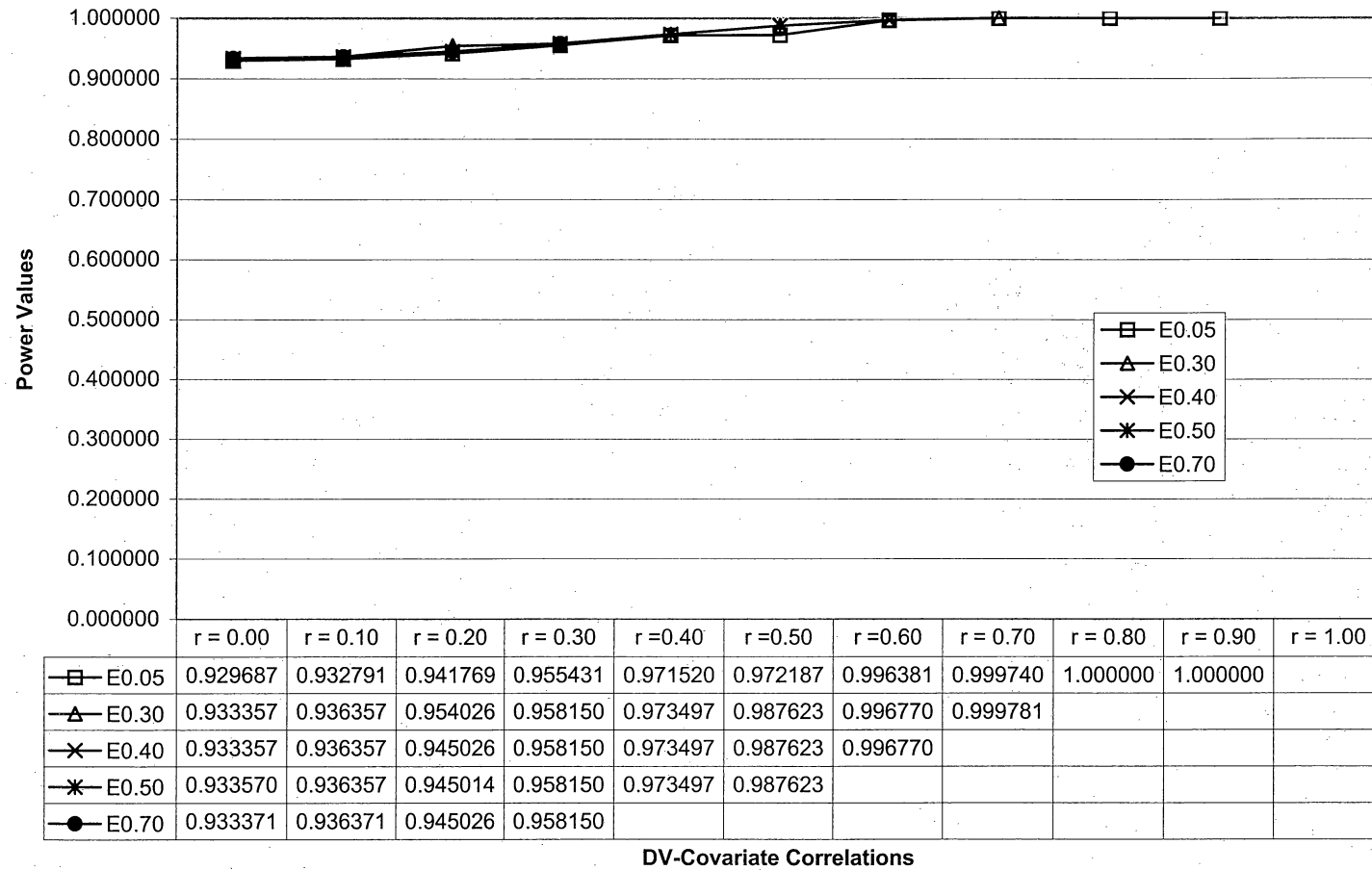


Figure 15

Power values across DV-Covariate correlations (r) and IV-covariate interactions (η^2) for $\alpha = .01$ and $n = 150$.

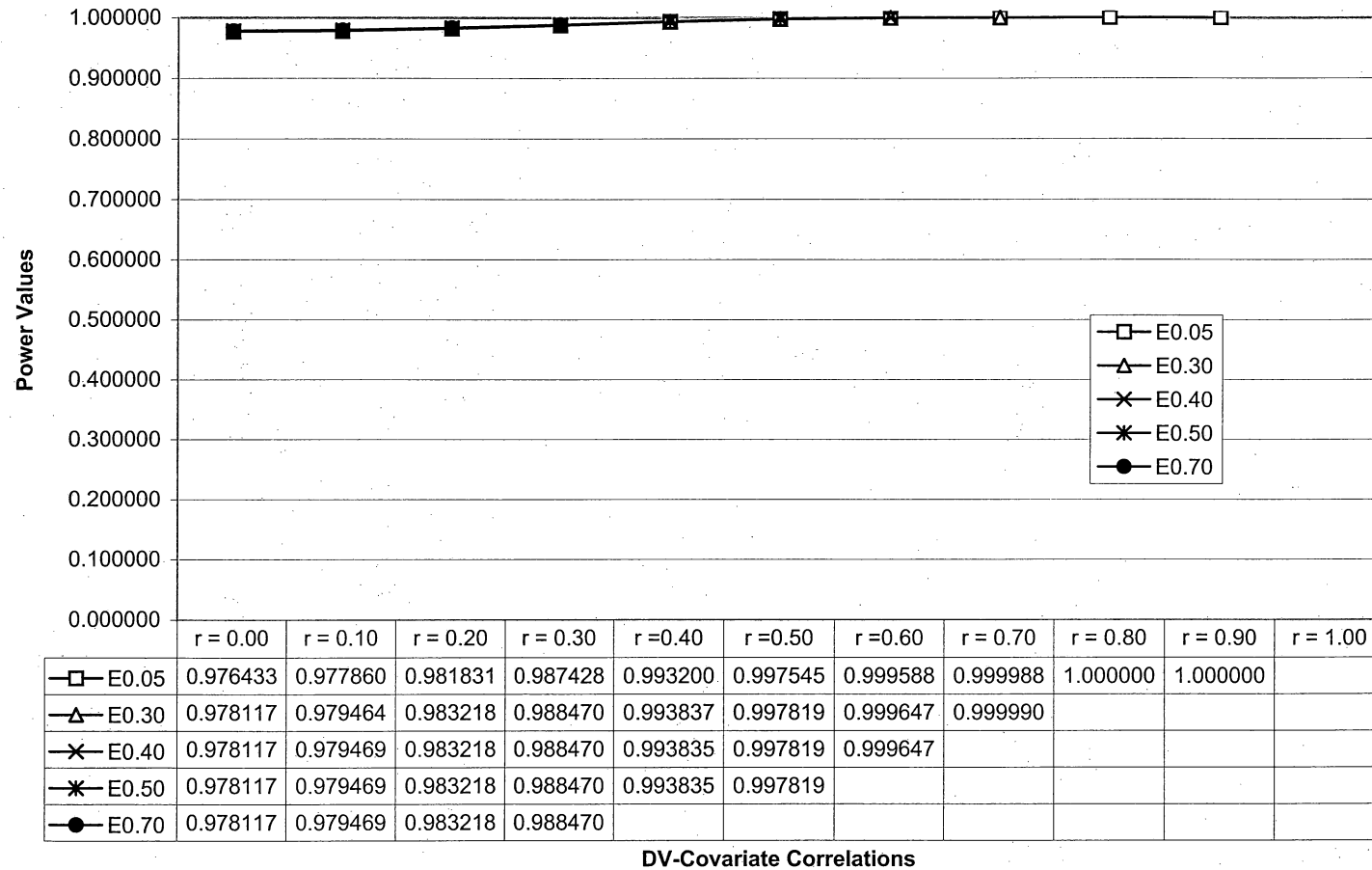


Figure 16

Power values across DV-Covariate correlations (\mathbf{r}) and IV-covariate interactions (η^2) for $\alpha = .05$ and $n = 30$.

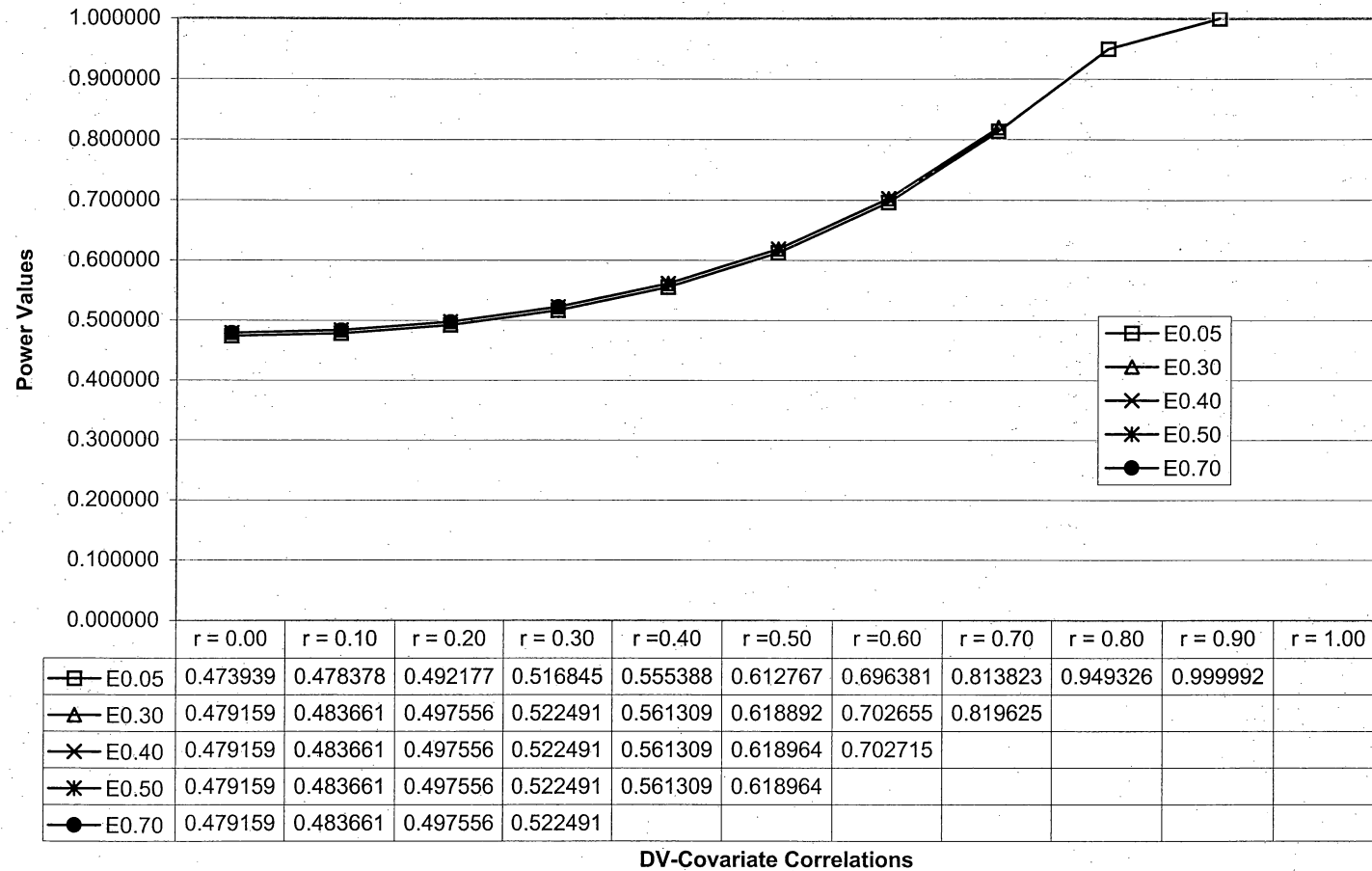


Figure 17

Power values across DV-Covariate correlations (r) and IV-covariate interactions (η^2) for $\alpha = .05$ and $n = 60$.

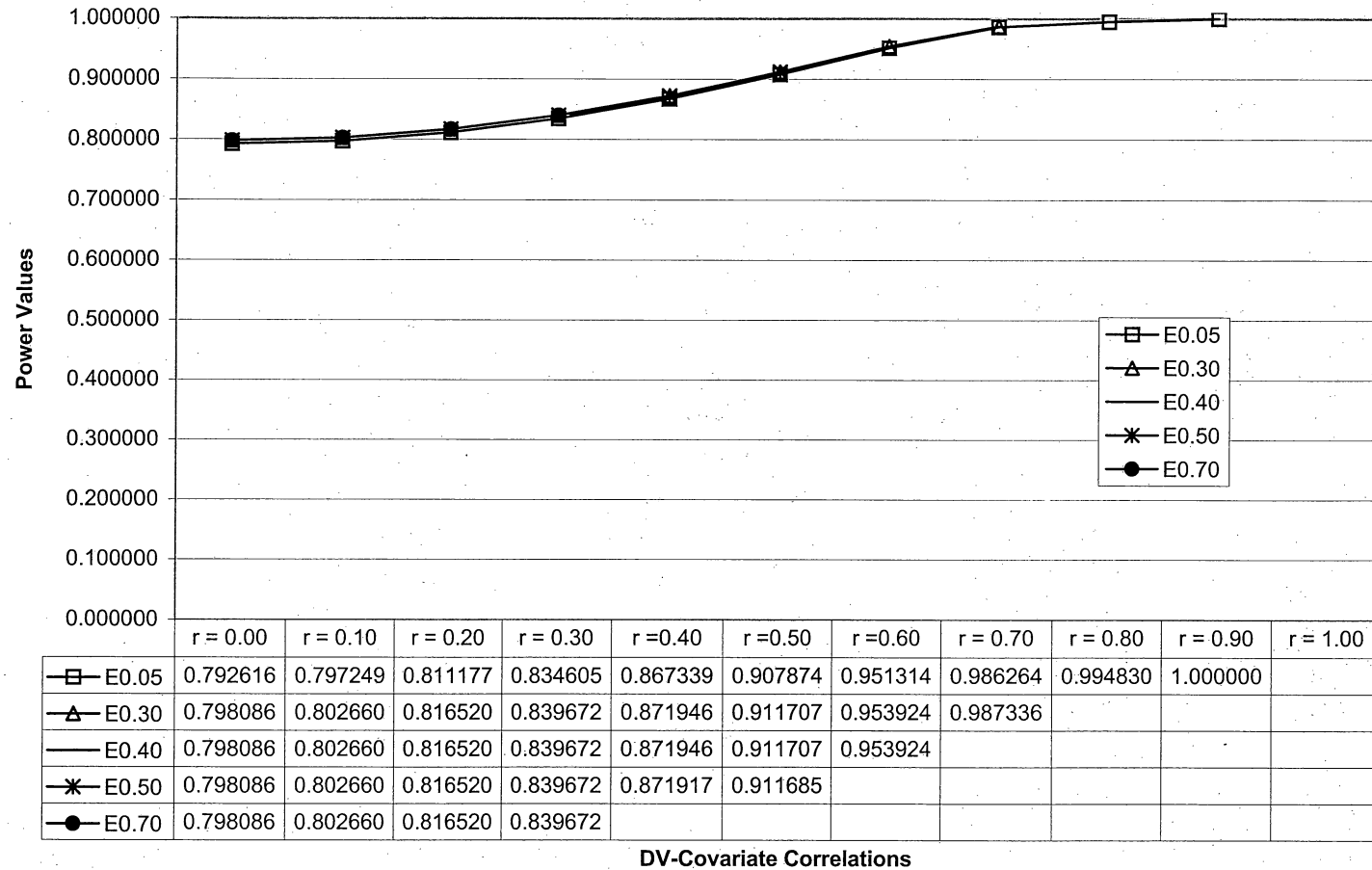


Figure 18

Power values across DV-Covariate correlations (r) and IV-covariate interactions (η^2) for $\alpha = .05$ and $n = 90$.

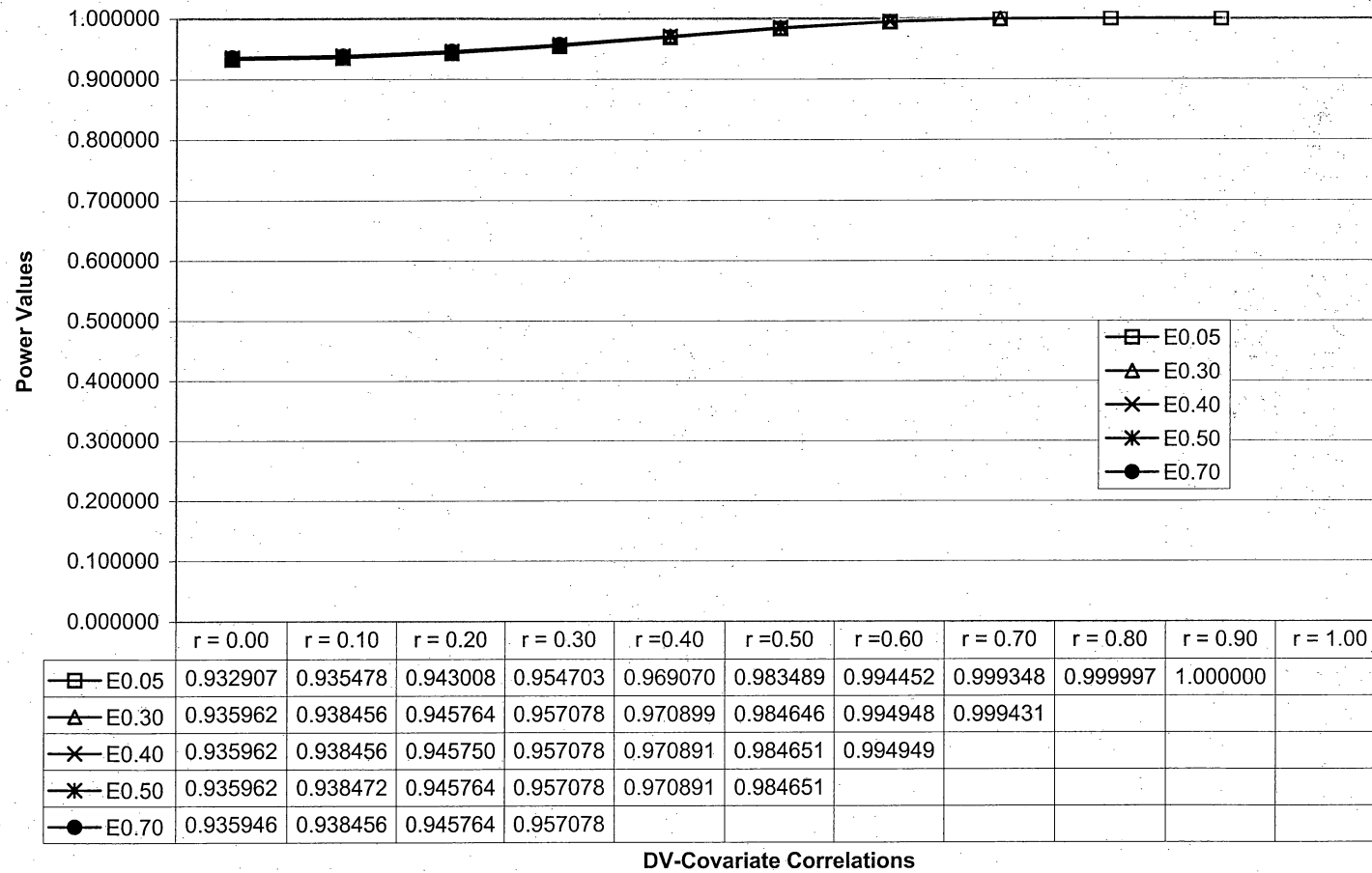


Figure 19

Power values across DV-Covariate correlations (r) and IV-covariate interactions (η^2) for $\alpha = .05$ and $n = 120$.

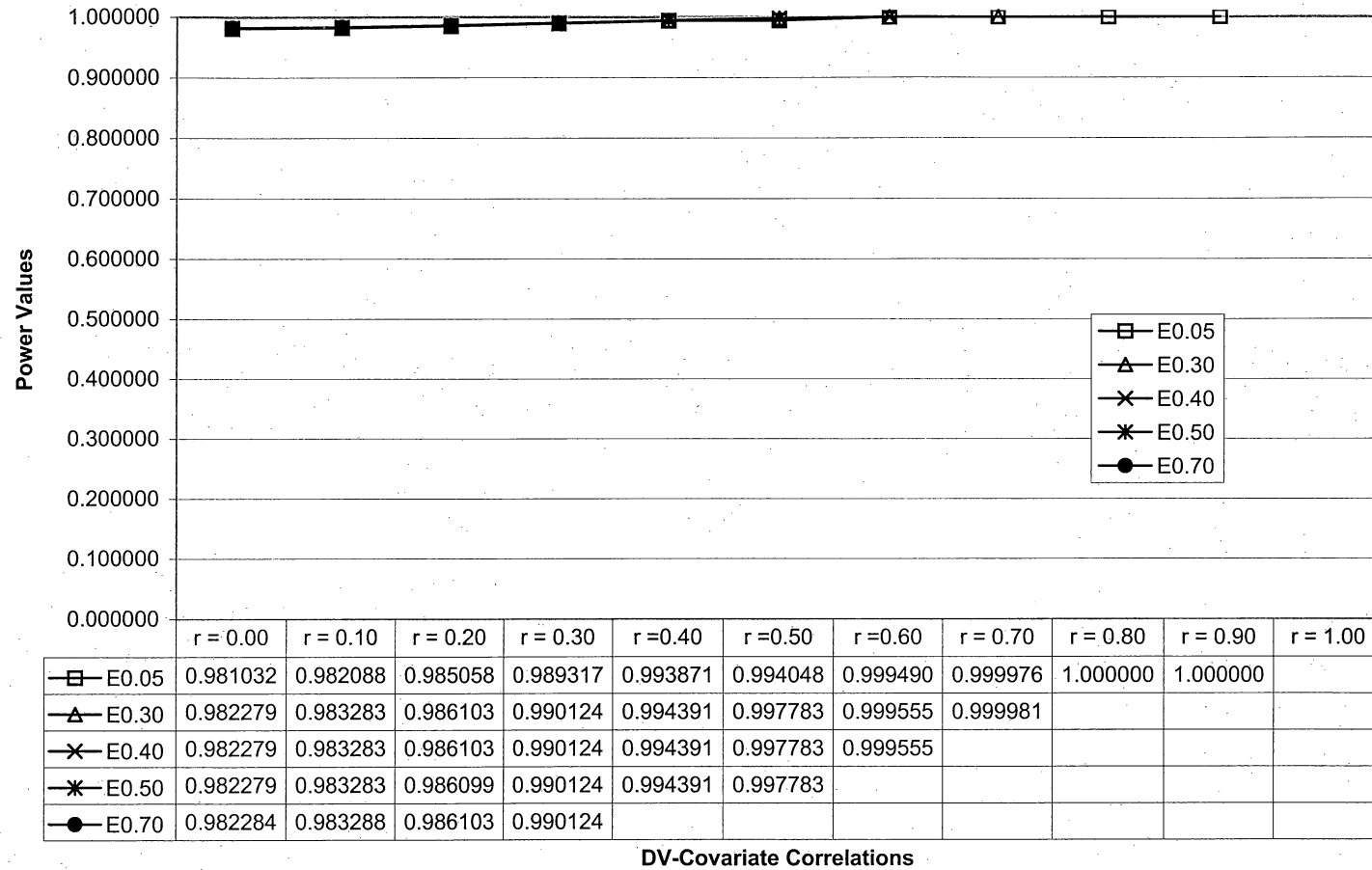
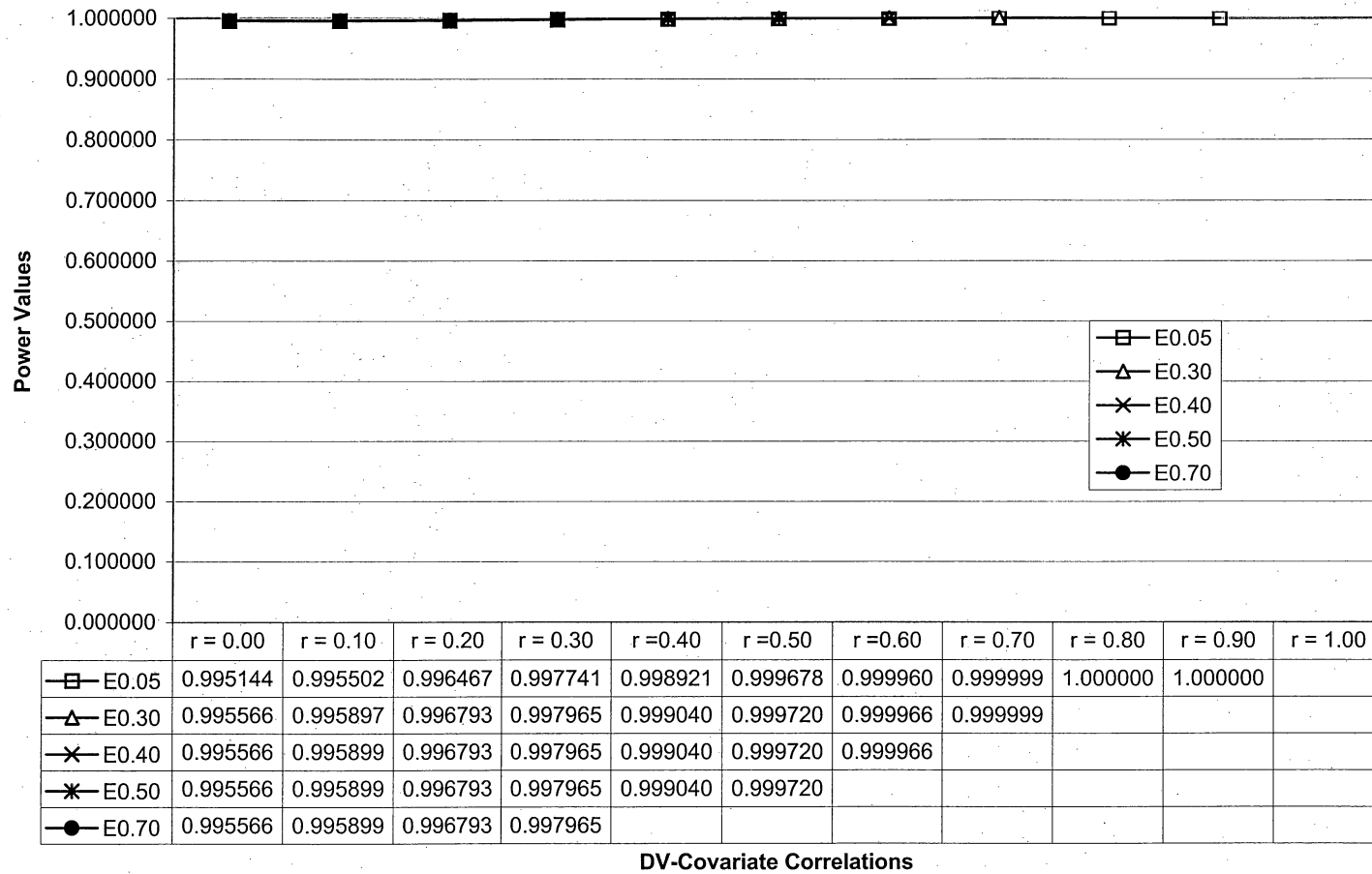


Figure 20

Power values across DV-Covariate correlations (r) and IV-covariate interactions (η^2) for $\alpha = .05$ and $n = 150$.



REFERENCES

- Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd Ed.). New Jersey: Lawrence Erlbaum Associates.
- Keppel, G. (1991). Design and analysis: A researcher's handbook (3rd Ed.). New Jersey: Prentice Hall.
- Kopriva, R.J. & Shaw, D.G. (1991). Power estimates: The effect of dependent variable reliability on the power of one-factor ANOVAs. Educational and Psychological Measurement, 51(3), 585-593.
- Kraemer, H.C. (1985). A strategy to teach the concept and application of power of statistical tests. Journal of Educational Statistics, 10, 173-195.
- Maxwell, S., Delaney, H., & Dill, C. (1984). Another look at ANCOVA versus blocking. Psychological Bulletin, 95(1), 136-147.
- Rogers, W.T. & Hopkins, K.D. (1988a). Power estimates in the presence of a covariate and measurement error. Educational and Psychological Measurement, 48, 647-656.
- Rogers, W.T. & Hopkins, K.D. (1988b). Quick power estimates incorporating the joint effects of measurement error and a covariate. Journal of Experimental Education, 57(1), 86-94.
- Tabachnick, B.G. & Fidell, L.S. (1996). Using Multivariate Statistics (3rd Ed.). New York: HarperCollins.

Ullman, J.B. (1997). Differential Power in Fixed Effects and Mean and Covariance Structural Models. Unpublished doctoral dissertation, University of California, Los Angeles.

Williams, R.H., Zimmerman, D.W., & Zumbo, B.D. (1995). Impact of measurement error on statistical power: Review of an old paradox. Journal of Experimental Education, 63(4), 363-370.